A Publication of the American Psychological Association

The Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

Articles	Page
The Use of the Thematic Apperception Test to Differentiate Normal from Abnormal Groups *Anne M. Ritter and Leonard D. Eron*	147
Parental "Identification" and "Tendency Toward Disorders" as Measured by the Minnesota Multiphasic Personality Inventory *Andrew L. Sopchak*	159
A Study of Anxiety and Learning *George Mandler and Seymour B. Sarason*	166
The Formation of Social Groups under Conditions of Success and Failure *J. C. Gilchrist*	174
Music and the General Activity of Apataetic Schizophrenics *Clyde G. Skelly and George M. Haslerud*	188
Index of Nailbiting in Children *Anthony J. Malone and Maury Massler*	198
Frustration, Satisfaction, and Other Factors Affecting the Attractiveness of Goal Objects *Robert J. Filer*	203
The Perceptual Consequences of Success and Failure *Leo Postman and Donald R. Brown*	213

Continued on inside front cover

Volume 47 APRIL 1952 Number 2

CONTENTS—*Continued*

ARTICLES PAGE

The Effect of an Audience on Behavior in a Choice Situation . . . 222
Seymour Wapner and Thelma G. Alper

Defense Against Ego-Threat in Memory and Perception 230
Charles W. Eriksen

The Interpersonal Process in Psychotherapy: Development of a Research Method 236
Allen T. Dittmann

CASE REPORTS

The Cultural Background of the Patient as Part of the Physician's Armamentarium 245
Harry Austryn Savitz, M.D.

CRITIQUE AND NOTES

Selective Perception in the Interpretation of Symbols 255
Harold H. Mosak and Frederick J. Todd

The Psychology of Religion: Some Suggested Areas of Research of Significance to Psychology 256
W. Edgar Gregory

Volunteer-Error in the Kinsey Study 259
A. H. Maslow and James M. Sakoda

Two Rorschach Scoring Techniques: The Inspection Technique and the Basic Rorschach 263
Anne Roe

A Note on the Validity of Sanford's Authoritarian-Equalitarian Scale 265
Joan Eager and M. Brewster Smith

REVIEWS 268

NEW BOOKS RECEIVED 281

Volume 47 APRIL 1952 Number 2

The Journal of Abnormal and Social Psychology

Editor: J. McV. Hunt, University of Illinois

Book Review Editor: Stanley G. Estes, Harvard University

Managing Editor: Lorraine Bouthilet

ASSOCIATE EDITORS

Gordon W. Allport
Harvard University

Hadley Cantril
Princeton University

Lee J. Cronbach
University of Illinois

John Dollard
Yale University

Leonard W. Doob
Yale University

Charles M. Harsh
Pomona College

William A. Hunt
Northwestern University

Arthur Jenness
Williams College

W. Ernst Kris
Yale University School of Medicine

Donald W. MacKinnon
University of California

Quinn McNemar
Stanford University

O. Hobart Mowrer
University of Illinois

Theodore M. Newcomb
University of Michigan

Charles E. Osgood
University of Illinois

Saul Rosenzweig
Washington University

Harold Seashore
Psychological Corporation

David Shakow
University of Illinois School of Medicine

Ross Stagner
University of Illinois

Elizabeth Collins Wesman
White Plains, N.Y.

Robert W. White
Harvard University

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

PUBLICATION OFFICE: 372 BROADWAY, ALBANY, N.Y., AND

1515 MASSACHUSETTS AVENUE, NW., WASHINGTON 5, D.C.

Entered as second-class matter, June 18, 1923, at the Post Office at Albany, New York, under the Act of February 28, 1925. Acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, authorized August 26, 1947.

Copyright 1952 by the American Psychological Association, Inc.

The JOURNAL is issued quarterly in January, April, July, and October. The yearly volume averages 544 pages. The subscription price per volume is $6.00; foreign, $6.50. Single issues are $1.75.

Original manuscripts and editorial communications should be sent to:

DR. J. MCV. HUNT
Department of Psychology
University of Illinois
Urbana, Ill.

Books for review should be sent to:

DR. STANLEY G. ESTES
Psychological Clinic
64 Plympton Street
Cambridge 38, Mass.

Five copies of the JOURNAL are supplied gratis to contributors of articles and book reviews. Authors may order reprints when returning proof. No reprints are supplied gratis.

Printed and mailed by:

THE BOYD PRINTING COMPANY, INC.
372 Broadway, Albany 7, N.Y.

Subscriptions and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1515 Massachusetts Avenue, NW.
Washington 5, D.C.

THE USE OF THE THEMATIC APPERCEPTION TEST TO DIFFERENTIATE NORMAL FROM ABNORMAL GROUPS[1]

ANNE M. RITTER AND LEONARD D. ERON
Yale University

In the fifteen years which have elapsed since Murray (8) introduced the Thematic Apperception Test, it has become one of the more widely used techniques in the clinical situation. The psychologist generally uses this examination to arrive at a formulation of the personality structure and dynamics of the individual tested. An objective scoring procedure, the examiner's hunches, or some combination of objective and intuitive interpretations are used to arrive at such evaluations.

Although in each succeeding year there has been an increasing number of research studies aimed at demonstrating the reliability and validity of the TAT (4, 9, 10, 11, 12, 14), the accumulation of normative data has been largely neglected until recently (2, 13). Despite the scarcity of normative data, various clinical investigators have stated that the presence of certain characteristics or "deviant" content in a TAT record provides a basis for differential diagnosis.

Rapaport (9), for example, has cited the occurrence of perceptual inaccuracies as a "sign" of schizophrenia. Other clinicians (4, 11, 14) have stated that themes of violence, hostility, death, destruction, guilt, and frustration are peculiar to particular clinical groups. The invalidity of these criteria, in view of the lack of norms, is pointed up in a recent normative study utilizing a normal control and several patient populations. The data of this study, described in detail elsewhere (2), indicate that perceptual distortions are no more representative of schizophrenic records than of any other groups of records; nor are any of the aggression, frustration, or guilt themes referred to above more characteristic of patient than of normal control records.

Generally, it was found (2) that the responses produced by each of two normal control groups, by nonhospitalized psychoneurotics, by hospitalized schizophrenics, and by a mixed neuropsychiatric hospital population to each card, when analyzed in terms of the formal structure and content of the stories, yielded only a few more statistically significant differences than could be expected by chance.

In view of the evidence which demonstrates that the fantasy content of various groups of normal and psychiatric patient populations is strikingly similar, it seems apparent that the mere presence of certain content or of a specific formal deviation cannot, as a rule, be used to differentiate various clinical groups. It was also shown that the physical structure of the individual card will, in itself, tend to arouse stories of a characteristic emotional tone, thematic content, etc. For at least some of the TAT pictures, then, this picture "pull" may be sufficiently strong to permit us, following an analysis of the productions of a normal control group, to designate certain themes, emotional tones, and outcomes of stories as the normally expected responses to these cards. Similarly, in terms of a formal analysis of TAT productions, the occurrence of perceptual deviations, of questionable outcomes, alternative themes, card rejections, and other interpretations contrary to instructions may be tabulated.

Such a tabulation of the frequency with which each individual fails to give expected responses and of the presence in each record of an excessive number of formal deviations of specific varieties may provide a statistically reliable and valid means of weeding out some proportion of individual records as abnormal. Thus, if it is possible to determine the frequency of deviations which only individuals in the abnormal groups exceed, this technique may be of value for broad diagnostic classification.

It is the aim of the present study to test

[1] The writers wish to thank Dr. Fred Sheffield, Yale University, for his assistance with and review of the statistical analysis and for his critical reading of the manuscript.

the validity of this assumption. The criteria available for such an analysis are at present very tentative and cannot be expected to separate out the bulk of each of the patient populations to be considered. It is believed by the writers, however, that there is sufficient normative data to allow a preliminary investigation along these lines.

The major assumptions made by the investigators, as previously stated, are that:

1. Certain content can be specified as the typical or expected response to various of the stimulus pictures.

2. Certain formal deviations can be specified as atypical and unacceptable interpretations of the instructions and/or of the stimulus pictures.

As regards content, the data will be analyzed in terms of the themes, the emotional tone, and outcome of each of the stories[2] by each subject. Formal deviations will be broken down in terms of perceptual distortions, questionable or inconclusive outcome of stories, and deviant interpretations other than perceptual distortions or questionable outcomes, e.g., card rejection, card description, denial of theme, autobiographical stories, etc.[3]

There is little evidence that there are any notable group differences between patient and normal populations. It is, however, expected that when deviant content and formal interpretations are tabulated for each individual in the patient and in the normal control populations, the median number of deviant interpretations per individual will be significantly lower in the normal than in the non-normal groups.

Similarly, it is expected that the range of deviations for individuals in the patient populations will be more extensive than in the normal control groups (i.e., some individuals in the patient populations will have many more deviations along some of the continua under consideration than will any individual in the normal control groups).

The data will be analyzed to determine the extent to which each of the content (*T, E, O*) and formal (*A, B, C*) deviations, individually and in various combinations, can be utilized to separate out individuals in the patient populations from the normal control groups.

[2] Theme, emotional tone, and outcome will be referred to as *T, E,* and *O* deviations respectively.

[3] Perceptual inaccuracies will be referred to as *A* deviations, formal deviations other than perceptual inaccuracies and questionable outcomes will be referred to as *B* deviations, and questionable outcomes will be referred to as *C* deviations.

METHOD AND PROCEDURE

Subjects

In the previously mentioned normative study (2),[4] the TAT was administered to 150 veterans in line with the instructions of Murray (8) except that all 20 cards were administered in one session. There were six groups with 25 *Ss* each. These groups, comparable in so far as age, educational level, intellectual level, and marital status are concerned, included the 50 *Ss* in the two unsystematically selected normal control groups (Groups I and II), 25 nonhospitalized psychoneurotics attending an outpatient V.A. clinic (Group III), 25 hospitalized psychoneurotics (Group IV), 25 hospitalized schizophrenics (Group V), and 25 individuals in the general neuropsychiatric hospital population (Group VI).[5]

Procedure

For purposes of this study, it is assumed that if a specific theme (or themes), emotional tone(s), or outcome to a TAT card is found in the stories related by 40 per cent of the normal control population, then this content may be considered an expected response. This is a considerably higher criterion than that used to specify the "popular" responses to the less structured Rorschach procedure (1, 5, 6). Since, however, the normative data available for this evaluation are derived from a relatively small population, rather stringent criteria should be applied. In any instance where no one theme is given by as much as 40 per cent of the normal control groups, it is assumed that no one theme can be expected to this card, and the theme variable is eliminated from consideration. Emotional tone and outcome are similarly treated.

The approximately 3,000 stories obtained from the six groups of subjects had, for the purposes of the original normative study, been carefully analyzed in terms of theme, emotional tone, outcome, perceptual distortions, levels of interpretation. These same data were now re-evaluated in line with the assumption that fantasy content which appears in 40 per cent or more of the normal control population's (20 *Ss*) stories to a given card may be considered a "popular" or "expected" response.

[4] In this monograph a detailed description of the make-up of each of the groups can be found; it also includes a detailed description of the method by which the check list of 125 themes was obtained and how the scales for rating emotional tone and outcome of story were used. A detailed account of the various levels of interpretation, i.e., formal deviations, is also included.

[5] The groups hereinafter referred to as Groups I, II, III, IV, V, VI are made up of the same individuals referred to as Groups C1, C2, NHPN, HPN, HS, and GNPH respectively in the earlier report.

The amount of normative data now available is insufficient to permit differential weighting of each of these variables for any card, alone or in combination. They are scored thus only in terms of presence or absence.

It should be pointed out that theme, emotional tone, and outcome are treated separately as well as jointly; i.e., if in response to a given card, one theme appears in 20 or more of the normal control records, while no emotional tone reaches this frequency criterion, then the data pertaining to the card will be treated in terms of theme, but not emotional tone. On the other hand, if the criterion for inclusion is met for emotional tone and for theme, then deviations from the expected theme and emotional tone will be tabulated for each individually and for theme and emotional tone combined.

The criterion was met on 14 cards for theme, 15 cards for emotional tone, 8 cards for outcome, and 11 cards for theme and emotional tone combined.

Tabulations of the deviations in the formal aspect of the productions were done separately. While each *T* (theme), *E* (emotional tone), or *O* (outcome) tally indicates the failure to give an "expected" response, any *A*, *B*, or *C* tally, as previously explained, indicates that the individual has somehow deviated from instructions. Whenever a perceptual inaccuracy was noted, an *A* deviation tally for that subject to the card involved was recorded. Although *B* and *C* deviations are both deviant responses in level of interpretation, they were separately recorded. *A*-type deviations were recorded for 17 cards, *B*-type deviations for 20 cards, and *C*-type deviations for 20 cards. When combined with theme (*T*) and emotional tone (*E*) deviations,[6] the following combinations are found possible:

ABCTE on cards 1, 4, 5, 6, 7, 8, 9, 10, 12, 13, 18, 20
ABC on cards 1, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 17, 18, 19, 20
BC on cards 1 through 20
A on cards 1, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 17, 18, 19, 20
B on cards 1 through 20
C on cards 1 through 20
TE on cards 2, 4, 5, 6, 7, 8, 9, 10, 12, 13, 18, 20
T on cards 1, 2, 4, 5, 6, 7, 8, 9, 10, 12, 13 18, 19, 20
E on cards 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 18, 20
O on cards 4, 5, 7, 8, 9, 10, 12, 17

The frequency with which each of these deviations (individually and in the combinations indicated) appeared in the records of each of 150 *Ss* was recorded.

Statistical Treatment of Data

Because the data to be analyzed consist of frequency data—specifically, frequency of deviations—the chi-square technique is employed to determine whether the non-normal groups (Groups III, IV, V, VI) differ significantly from the two normal groups.

The total number of formal and content deviations in each group is presented in Table 1. In Table 2, the range, score closest to the median, and quartile deviation of *T, E, A, B, C* deviations individually and in the various combinations for each group are presented. The data for Groups I and II are presented individually. For each statistical analysis they are combined and treated as one group with an *N* of 50.[7] In comparing the normal control groups with Groups III, IV, V, VI, a dividing point as near as possible to the median[8] was determined by throwing into one distribution the data for the 50 normals (Groups I and II) and the 25 *Ss* in the particular non-normal group under consideration. The proportion of *Ss* above the "median" in the normal group and in the non-normal group was then determined, and the reliability of the difference between the standard error of these proportions was computed.[9]

A similar procedure was followed in another set of comparisons of the same distributions. In this second analysis, the upper ends (high frequency of deviations) of the distributions of Groups I and II were used as cut-off points. It was then determined whether the number of individuals in each of the other groups who fall at or beyond this point is significantly greater than in Groups I and II. The use of a cut-off close to the median reveals whether there are significant differences at the middle of the distribution of deviations. It may, however, mask real differences at the extremes. The consideration of high frequency of deviations is particularly important in considering the data of the nonhospitalized psychoneurotics. The a priori assumption can be made that individuals receiving outpatient psychiatric treatment are significantly more "maladjusted" and "deviant" in their behavior, mental content, inter-

[6] Omission of *O* deviations from these combinations will be explained later.

[7] The distribution of each final deviation between the two normal control groups was first subjected to a chi-square analysis to determine whether these two groups can be treated as one population. Of the 20 chi-square values, 19 of the differences were well within the limits of chance expectation at the .05 level of confidence.

[8] This point will be referred to as the "median."

[9] The formula used to compute the reliability of the difference between proportions above and below the dividing point closest to the median was:

$$\sigma_{p_1-p_2}=\frac{p_1-p_2}{\sqrt{(PQ)\left(\frac{1}{N_1}+\frac{1}{N_2}\right)}}$$

where p_1=proportion of *Ss* above the "median" in Group 1
p_2=proportion of *Ss* above the "median" in Group 2
$P=\frac{F_1+F_2}{N_1+N_2}$
$Q=1-P$
N=total number of *Ss* in groups under comparison

personal relations, etc., than are the bulk of the normal control population. On the other hand, it is likely that there are individuals in a "normal control group" who may be "neurotic" and more in need of treatment than some of the people who are receiving outpatient psychiatric care. In other words, it can be expected that within the normal control group a few individuals may have a large number of deviations. And within each patient-population group there will be individuals who give many or most of the "expected" responses and few of the formal deviations, while there will be others who reverse this procedure. As we proceed from the nonhospitalized psychoneurotics to the various hospitalized groups, the records will be likely to contain fewer expected responses and more formal deviations. However, regardless of the number of individuals in each group whose productions do not contain deviations in excess of the median, some measure is needed of the extent to which each group contains a significant proportion of individuals whose deviations do exceed those of the normal group. For this analysis, the high frequency cut-off point was taken at that frequency which included two individuals of the 50 in the normal group. The reliability of the difference in the proportion of individuals in each of the non-normal groups whose frequency of deviations reach or exceed that cut-off point was computed. In each case of computing the reliability of difference in the proportions, the expected frequency is based not on the performance of the normal group alone; the combined frequency of the normal groups and the non-normal group is used to determine the expected frequency.

TABLE 1

TOTAL NUMBER OF DEVIATIONS OF EACH VARIETY (PERCEPTUAL DISTORTION, QUESTIONABLE OUTCOME, LEVEL OF INTERPRETATION, THEME, EMOTIONAL TONE), INDIVIDUALLY AND COMBINED, BY THE 25 SUBJECTS IN EACH GROUP

DEVIATION	FREQUENCY IN EACH GROUP					
	I	II	III	IV	V	VI
A, Perceptual Distortion	53	39	57	55	69	87
B, Levels of Interpretation	63	62	85	120	149	116
C, Questionable Outcome	24	26	68	104	74	76
Total of *A*, *B*, and *C*	140	127	210	279	292	279
T, Themes	122	126	142	147	157	156
E, Emotional Tone	127	117	143	178	176	185
Total of *T* and *E*	249	243	285	325	333	341
Total of *A*, *B*, *C*, *T*, *E*	389	370	495	604	625	620
(Outcome)	(103)	(102)	(103)	(104)	(106)	(105)

RESULTS

Comparison of Median Score Deviations of Normal and Non-Normal Groups

Themes, emotional tone, and outcome deviations. In contrast to the 13 pictures which met the criterion for inclusion in regard to theme and the 16 in regard to emotional tone, similar outcomes were given by 40 per cent of the normal group to only eight of the pictures. And no one of these differentiates any group from any other. The total number (Table 1) of *O* deviations and their distribution in Groups III, IV, V, VI differed from those given by Groups I and II no more than would be expected by chance in any instance. *O* deviations, thus, provide no means by which abnormal records can be differentiated from normal records. Nor are any differences in range, "median," or quartile deviation found between any groups (Table 1). It is because of the complete lack of differentiation with this criterion that "outcome" has not been used when the differentiating effectiveness of combinations of the other deviations is tested. Inasmuch as outcomes are more an individual projection than any of the other variables (3), it is not surprising that no broad group differences obtain.

Both *T* and *E* deviations, however, whether considered separately or in combination (Tables 1, 2, 3), yield distributions in which a significant proportion of non-normal *S*s deviate reliably more from the median number of deviations than do normal *S*s. With the exception of Group III, whose distribution of *T* deviations just misses being significantly different at the .05 level of confi-

dence (p=approximately .07), all other p-values are approximately .01 or better.

Perceptual distortions, questionable outcome, and other deviations in level of interpretation. The rangė, "median," and quartile deviation of perceptual inaccuracies (*A* deviations), of deviations in level of interpretation other than questionable outcome (*B* deviations), and of questionable outcomes (*C* deviations), of *B* and *C* deviations combined, and all *A, B, C* deviations combined are presented in Table 2. From an inspection of these tables, two things are apparent: first, the range, "median," and quartile deviations in the normal groups are lower than those of Groups IV, V, VI in every instance; although the Group III deviations are somewhat higher along some continua, they differ less consistently from the normal groups than do the other patient groups. Second, the distribution of deviations for the normal groups is, at least for the *B* and *C* deviations individually and combined, much narrower than it is for the non-normal groups. The impression gathered from such inspection is, in most instances, found to hold up when the data are subjected to more detailed statistical analysis. This is evident from the chi-square values obtained in testing the hypothesis that Groups I and II combined vs. Group IV, vs. Group V, and vs. Group VI are homogeneous populations in so far as the frequency distribution of *B* and *C* deviations within each of these combinations of groups is concerned. In each of these six instances the obtained p-values which range from .05 to less than .01 (Table 3) indicate that they are not homogeneous populations, but rather that there are significantly more individuals in the normal group whose scores are low (below the median) than in the other groups.

The null hypothesis, when applied to *A* deviations, however, is not rejected in three of the four tests. It is only when the normal groups are compared with the mixed neuro-

TABLE 2

RANGE, "MEDIAN," AND QUARTILE DEVIATION OF EACH GROUP INDIVIDUALLY AND THE TWO NORMAL GROUPS COMBINED FOR EACH OF THE DEVIATIONS

VARIABLE	STAT. MEASURE	I	II	I–II COMBINED	III	IV	V	VI
T	Range	2–9	3–8	2–9	3–10	3–9	4–10	3–10
	"Median"	5	5	5	6	6	6	6
	Q	0.94	0.96	0.97	1.30	1.30	1.14	0.94
E	Range	2–9	2–8	2–9	2–10	3–11	3–12	5–11
	"Median"	5	5	5	6	7	7	8
	Q	1.17	1.40	1.24	1.68	1.22	1.67	1.14
TE	Range	6–16	6–14	6–16	5–17	6–18	8–22	10–18
	"Median"	10	10	10	11	13	13	13
	Q	1.92	1.56	1.72	2.51	1.86	2.06	1.47
A	Range	0–4	0–4	0–4	0–6	0–5	0–7	0–6
	"Median"	2	1	2	2	2	2	4
	Q	0.98	0.60	0.80	0.84	0.89	1.18	1.06
B	Range	0–7	0–8	0–8	0–11	0–10	0–21	0–17
	"Median"	2	2	2	3	4	3	3
	Q	1.20	1.23	1.22	1.36	2.15	3.25	2.63
C	Range	0–3	0–4	0–4	0–11	0–12	0–13	0–13
	"Median"	1	0	0	1	3	2	3
	Q	0.60	0.79	0.68	2.38	2.50	1.54	2.12
BC	Range	0–8	0–10	0–10	0–18	0–22	0–24	0–30
	"Median"	3	3	3	4	8	6	6
	Q	1.64	2.02	1.84	4.10	3.72	3.16	4.06
ABC	Range	1–11	2–13	1–13	2–20	1–25	2–38	1–34
	"Median"	5	3	4	5	9	9	9
	Q	2.27	2.43	2.48	5.10	4.34	4.06	4.16
ABCTE	Range	8–22	8–24	8–24	10–31	11–41	10–49	14–47
	"Median"	15	14	14	18	22	23	23
	Q	1.92	2.34	2.15	3.00	3.60	3.82	3.66

TABLE 3

RELIABILITY OF DIFFERENCE BETWEEN THE PROPORTION OF SUBJECTS AT A DIVIDING POINT CLOSE TO THE MEDIAN BETWEEN EACH NORMAL GROUP, BETWEEN NONHOSPITALIZED AND HOSPITALIZED PSYCHONEUROTICS, BETWEEN THE TWO NORMAL GROUPS COMBINED AND EACH OF THE NON-NORMAL GROUPS
(Chi-square and p-values significant at or beyond the .10 level are indicated.)

VARIABLE	BETWEEN I AND II *			BETWEEN III AND IV			BETWEEN (I AND II) AND											
							III			IV			V			VI		
	σD_p	χ^2	p	σD_p	χ^2	p	σD_p	χ^2	p	σD_p	χ^2	p	σD_p	χ^2	p	σD_p	χ^2	p
T	0.90	0.81		0.58	0.33		1.82	3.33	.07	2.47	6.09	.02	2.79	7.77	.01	3.11	9.12	<.01
E	0.00	0.00		1.70	2.89	.10	1.65	2.73		3.92	15.38	<.01	2.94	8.65	.01	4.93	24.35	<.01
TE	0.86	0.74		1.41	1.99		0.82	0.67		3.80	14.42	<.01	3.45	11.93	<.01	3.45	11.93	<.01
A	1.74	3.02	.10	0.87	0.76		1.01	1.00		1.34	1.80		1.34	1.80		3.81	14.54	<.01
B	0.86	0.74		2.55	6.50	.02	0.82	0.67		3.45	11.93	<.01	2.12	4.51	.05	2.12	4.51	.05
C	0.85	0.72		1.70	2.89	.10	1.15	1.32		3.45	11.93	<.01	2.82	7.95	.01	1.15	1.32	
BC	0.00	0.00		1.98	3.92	.05	0.99	0.97		4.09	16.73	<.01	3.61	13.02	<.01	2.61	6.83	.01
ABC	1.15	1.32		1.41	1.99		1.80	3.23	.07	3.60	12.96	<.01	3.28	9.55	.01	3.92	15.37	<.01
ABCTE	1.18	1.39		1.99	3.96	.05	1.80	3.23	.07	4.90	24.04	<.01	4.25	18.03	<.01	4.57	20.91	<.01

* In terms of the assumptions underlying these analyses, significant differences between Groups I and II are not expected.

psychiatric population which makes up Group VI that significant differences emerge. Two possible reasons why four of six tests are not significant are suggested. It is possible that both very few perceptual distortions and very many perceptual distortions are as likely to occur in a neurotic (both outpatient and hospitalized) and a schizophrenic population as in a normal population. On the other hand, it is possible that the comparatively high number of *A* deviations in one of the two normal groups is not representative of the frequency of perceptual deviations characteristic of normal group behavior. From the currently available data, it cannot be determined which is the more accurate reason. Replication with comparable groups would be a necessary prerequisite to determine whether perceptual distortions in general or particular perceptual inaccuracies should be included among these criteria.

It should be pointed out that in none of these four comparisons was Group III found to differ significantly from Groups I and II. Although this is in line with Eron's finding (2) that the major differences are found between the hospitalized and nonhospitalized groups and that the fantasy productions of a nonhospitalized psychoneurotic group differ little from those of a normal population, inspection of the frequency distributions [10] suggested greater variability in Group III than in Groups I and II and, as will be demonstrated in a later section, this hypothesis appears valid.

Formal and content deviations combined. When both formal and content deviations are combined (*ABCTE*), then the difference between the groups is greatly increased. From the frequency distributions it was found that 76 per cent of Groups I and II have less than 17 deviations, but only 40 per cent of Group III, 12 per cent of Group IV, 20 per cent of Group V, and 16 per cent of Group VI have equally few deviations.

In each group, then, there are some individuals whose TAT productions include the expected content and exclude deviations; there are many less such individuals among the nonhospitalized psychoneurotics than among the normal control populations. On the other hand, the difference between the number of individuals in Group IV, in Group V, and in Group VI, as compared with those in Groups I and II, who exclude the expected material and include the deviant in their productions is a difference which could have occurred on a chance basis less than once in a thousand times (Table 3).

In comparing Groups I and II with Group III, the level of confidence of the difference is .07. It would not be expected that individuals undergoing psychiatric treatment in an outpatient clinic would include as many individuals who produce a large number of deviations as are found in the inpatient population. In terms of both the range of deviations from expectation and "median" for the group, these nonhospitalized psychoneurotics fall about midway between the normal control group and the hospitalized groups. This is apparent from an inspection of the total number of deviations (Table 1) from the range, "median," and quartile deviation data in Table 2. The second and third columns of Table 3 indicate that Group III differs approximately as much from the normal groups as from Group IV.

Comparison Between Number of Individuals in Normal and Non-Normal Groups Having High Frequency of Deviations

It has been pointed out that for all of the variables under consideration except outcome of stories, the frequency distribution of deviations per individual appears to show greater variability for each of the non-normal groups than for Groups I and II.

From an inspection of frequency distributions and of the data in Table 2, it can be determined that in Groups I and II, taken individually or in combination, Q_3 and Q_1 are approximately equal. This indicates that the distributions are essentially symmetrical. In the case of Group III, while Q_1 is in most instances as large or larger than Q_1 for Groups I and II, Q_3 is considerably higher. Thus, from the low-frequency deviations from the "median," Groups I and II do not differ significantly from Group III. Beyond the central cut-off point, however, the distributions of Group III deviations are considerably extended, i.e., positively skewed. This skewness is more marked for Groups IV, V, and VI than it is for Group III.

[10] The frequency distributions of each kind of deviation in each group are not presented because of space limitations.

To determine the validity of this judgment from inspection that significantly more individuals in the non-normal groups have a higher number of deviations than do the individuals in Groups I and II, the high frequency cut-off points previously described were used to determine whether the positive skewness of the non-normal group distributions involves significantly more deviations than are found in the normal groups. The frequency distributions indicate that in Groups I and II combined two individuals had as many as 14 to 16 deviations from the expected *TE* combined and no subject had more than 16. In the other groups, the number of individuals having 14 to 16 or more *TE* deviations is 8 in Group III, and 12 each in Groups IV, V, and VI. After each of these frequencies is translated into proportions for each group, the reliability of the difference between these proportions is found to be significant at the .001 level of confidence as between the normal groups and each of the other groups. Of the 36 reliability of differences and chi-squares computed, 30 are significant at or beyond the .05 level of confidence.

The nonhospitalized psychoneurotic group, which demonstrated no differences significant at or beyond the .05 level of confidence in the central portion of the distribution, was found to differ significantly from the normal groups in six of the nine combinations of deviations considered. From this we may presume that there is a real difference between the number of individuals in the non-normal groups whose deviations will exceed a given point and the number in the normal group who exceed that point.

No significant differences between the two psychoneurotic groups are found. In comparing high frequency deviations of Group IV with those of Groups I and II, 7 of the 9 tests yield differences significant at or beyond the .05 level of confidence; between the normal controls and Group V and the normal controls and Group VI, the number of tests yielding significant differences are 9 and 8 out of 9 respectively. Six of the 9 tests reveal significantly more individuals to deviate more frequently in Group III than in I and II.

The only tests which fail to yield significant differences in any instance are some of those in which deviations are considered individually, e.g., *T, E, B* alone. When these deviations are combined, so that, for example, both questionable outcome and other level discrepancies are treated together, when both *B* and *C* deviations are combined with *A* deviations, when *T* and *E* are combined, or when all are treated together (*ABCTE*), then the level of significance between the normal and each of the non-normal groups, including Group III, is very high.

DISCUSSION

From these data it can be inferred that, even though with the individual deviation criteria a significant number of individuals in the hospitalized group can be separated out from the total population, the single criterion is insufficient to separate out more than a few of the nonhospitalized psychoneurotics. This might be expected from the findings earlier presented by Eron (2) to the effect that all themes, various perceptual inaccuracies, and level deviations are found in all groups, and that in terms of such over-all productions significant differences between the groups are found when they are divided into hospitalized vs. nonhospitalized. Despite the fact that all varieties of deviation are found in all groups and that some individuals in all of the hospitalized groups deviate only occasionally, the distribution of deviations for each of the hospitalized groups appears to be significantly different from that of the normal group. This can be inferred from the fact that, in respect to most of the deviations considered, the hospitalized groups were usually found to differ significantly from the normal groups in terms of both "median" and high deviation scores. In comparing the "median" scores of Groups I and II with Group III, some differences are indicated. These differences, though greater than those between the two normal control groups, were not for the most part significant. Thus the productions of at least 50 per cent of the nonhospitalized neurotics are in line with the productions expected of the normal population. On the other hand, when the tail end of the distribution is considered, a significant proportion of the subjects in Group III gives reliably more deviations than do

TABLE 4

RELIABILITY OF THE DIFFERENCE BETWEEN THE PROPORTION OF SUBJECTS IN EACH NON-NORMAL GROUP AND THE PROPORTION IN THE NORMAL GROUPS WHO EXCEED A HIGH FREQUENCY CUT-OFF POINT FOR EACH OF THE DEVIATIONS SCORED AND RELIABILITY OF THESE DIFFERENCES BETWEEN THE NONHOSPITALIZED AND HOSPITALIZED PSYCHONEUROTICS

(Chi-squares and *p*-values at or beyond the .10 level of significance are reported.)

VARIABLE	BETWEEN III AND IV			BETWEEN (I AND II) AND											
				III			IV			V			VI		
	σD_P	χ^2	p	σD_P	χ^2	p	σD_P	χ^2	p	σD_P	χ^2	p	σD_P	χ^2	p
T	0.00	0.00		1.81	3.28	.10	1.81	3.28	.10	2.64	6.97	.01	1.21	1.46	
E	1.77	3.13	.10	2.26	5.11	.05	4.29	18.37	<.01	3.98	15.84	<.01	4.88	23.81	<.01
TE	1.15	1.32		3.36	11.28	<.01	4.61	21.25	<.01	4.61	21.25	<.01	4.61	21.25	<.01
A	0.41	0.17		0.75	0.57		0.26	0.07		2.00	4.01	.05	4.91	24.11	<.01
B	0.66	0.44		1.51	2.28		2.31	5.34	.05	3.35	11.22	<.01	2.31	5.34	.05
C	1.18	1.39		3.02	9.12	.01	4.31	18.58	<.01	3.02	9.12	.01	4.01	16.08	<.01
BC	0.89	0.78		3.36	11.28	<.01	4.31	18.58	<.01	3.69	13.62	<.01	3.02	9.09	.01
ABC	0.00	0.00		3.69	13.62	<.01	3.69	13.62	<.01	3.69	13.62	<.01	3.02	9.09	.01
ABCTE	1.01	1.02		3.36	11.28	<.01	5.46	29.81	<.01	5.46	29.81	<.01	5.74	32.95	<.01

members of the normal group. When the variability of the groups is taken into account, then the groups more readily divide themselves into five populations: one normal group and four non-normal groups.

Formal and content deviations are both more significant when considered in combination. This demonstrates that individuals in any group—normal or otherwise—do not give all the expected responses and frequently include responses deviant from instructions. In the normal groups, however, such deviations occur more or less at random throughout the cards so that, although an expected theme may not be given, the emotional tone of the story which is given is in line with the "picture pull" expectations of the card. Such a subject's record then would yield a *T* deviation for that card, but not an *E* deviation, and thus not a *TE* deviation. In each of the non-normal groups, however, to cards to which the subject fails to give the expected theme, there are significantly more individuals who also fail to give the expected tone. These differences are more sharply defined when all the deviations—*ABCTE*—are combined. Thus, it would seem that the appropriate combination of deviations from expected responses can be used to separate out certain records as abnormal. The criteria here presented are tentative and should be replicated before they are used in the clinical situation.

These results, however, do have some bearing on interpretations of individual protocols. It would seem that a sentence-by-sentence, word-by-word dissection of a subject's productions (11, 12), giving equivalent importance to every utterance without regard to the properties of the fantasy-evoking stimuli and/or their valence for a large number of individuals, is not a valid procedure. Fantasy is significant for diagnostic purposes only when it deviates from the norm either in frequency or kind. In interpreting fantasy productions these are the basic data—the frequency and unusualness of specific fantasy content—which must be specified before any interpretive significance can be placed on the thematic material obtained. How any analyst will then use these data to interpret personality structure and content will depend on his own theoretical orientation and on the amount of behavioral data obtained from other sources which are available to him and in the light of which he can ascribe meaning to the fantasy material.

Summary and Conclusions

TAT records of 150 *Ss* approximately matched for age, education, intellectual level, and marital status were used in this study. The *Ss* were divided into six groups of 25 each: two normal control groups, one group of outpatient psychoneurotics, one group of hospitalized psychoneurotics, one group of hospitalized schizophrenics, and one group consisting of a mixed neuropsychiatric population.

The assumption is made that if 40 per cent of the normal control groups respond to a picture with the same theme, emotional tone, and outcome these may be deemed "popular," and so, "expected" responses. The failure to give such "expected" responses is considered deviant. Similarly, responses which include perceptual inaccuracies, or an approach to the card in which the instructions are not followed (e.g., rejections, theme denial, alternate themes or outcomes, and other deviant levels of interpretations) are all considered to be deviant responses.

The records of each of the 150 *Ss* were analyzed to determine the frequency of occurrence of each such deviation, individually and in various combinations, to test the hypothesis that in each of the non-normal groups more individuals will deviate more frequently than do individuals in the normal groups.

The frequency with which any such deviation was found was tabulated for each individual on each card. The reliability of the difference between the proportion of subjects in the normal control and each of the other groups above and below the "median" number of deviations, the χ^2 and *p*-values of these distributions of deviations were computed for *T, E, A, B, C, TE, BC, ABC,* and *ABCTE* deviations. A similar analysis using a more extreme cut-off point (high frequency) in each distribution was also done.

The findings are:

1. Using the "median" as cut-off point, in 24 of the 36 tests the σ_{D_p} and χ^2 values indicated that significantly more individuals in the non-normal groups deviated from

expectation at a reliably higher frequency ($p=.05$ or better) than did the individuals in the normal control group.

2. Both theme and emotional tone deviations, individually and combined, in the hospitalized groups were significantly higher than those in the normal groups at or beyond the .01 level of confidence.

3. In respect to the outcome variable, no differences in excess of chance between the normal and patient groups were found and this variable was discarded.

4. The frequency of occurrence of questionable outcome and other deviations in level of interpretation were found to differentiate reliably between the normal groups and each of the hospitalized groups. The usefulness of perceptual inaccuracies is at present questionable.

5. The various formal and content deviations when combined (*TE, ABC, ABCTE*) more clearly and reliably differentiate more individuals in the non-normal groups than does any one deviation by itself.

6. Using the tail end of the distribution, 30 of the 36 χ^2 values indicated that there was a significantly greater range in the deviations of each of the non-normal than the normal groups.

7. When all deviations combined (*ABCTE*) are considered, each of the non-normal groups differs significantly from the normal groups at the .001 level of confidence.

8. In each group there are some individuals who do not deviate from expectation with any greater frequency than do the members of the normal group. Within each non-normal group, however, there is a significant number of individuals whose frequency of deviation from expected responses is reliably greater than in the normal group. Among the nonhospitalized psychoneurotics this is found for fewer individuals and with a lower frequency than in the other non-normal groups. Among the nonhospitalized psychoneurotics, for at least 50 per cent of the group the distribution of deviations does not differ from that of the normal group beyond chance expectation. The variability of the former, however, is greater, and so there are a significant number of individuals who have reliably more deviations than are found among the normals. Significant differences, thus, obtain at the high-frequency deviation extreme of the distribution even where such differences are not significant in the middle of the distribution.

9. No differences in excess of chance expectation were found between the two normal groups. The nonhospitalized psychoneurotic group falls between the normal and hospitalized psychoneurotics and differs approximately as much from the former as from the latter. The hospitalized psychoneurotics, schizophrenics, and mixed hospitalized neuropsychiatric populations fairly consistently deviate significantly from the normal control groups.

10. It is concluded that this method of analysis appears worthy of further investigation. The possibility of its use to weed out some proportion of deviant individuals and, eventually, as a tool in diagnostic classification is suggested. In the interpretation of a given protocol it is felt that this study offers further evidence that fantasy material is important chiefly when it deviates from the expected response. The need for replication and more extensive normative data are pointed out.

11. An item analysis of each variable, individually and in various combinations, in order to separate the best differentiating items and so provide a brief and easily used scoring procedure is proposed as a further research problem.

REFERENCES

1. BECK, S. J. *Rorschach's test.* Vol. I: *Basic processes.* New York: Grune & Stratton, 1944.
2. ERON, L. D. A normative study of the Thematic Apperception Test. *Psychol. Monogr.*, 1950, **64**, Whole No. 315.
3. GARFIELD, S. L., & ERON, L. D. Interpreting mood and activity in Thematic Apperception Test stories. *J. abnorm. soc. Psychol.*, 1948, **43**, 338–345.
4. HENRY, W. E. The Thematic Apperception Technique in the study of culture-personality relations. *Genet. Psychol. Monogr.*, 1947, **35**, 3–315.
5. HERTZ, M. R. *Frequency tables to be used in scoring responses to the Rorschach Ink-Blot Test.* Cleveland: Western Reserve Univ. Press, 1946.
6. KLOPFER, B., & KELLY, D. McG. *The Rorschach technique.* New York: World Book Co., 1942.
7. MCNEMAR, Q. *Psychological statistics.* New York: Wiley, 1949.

8. MURRAY, H. A. Techniques for a systematic investigation of phantasy. *J. Psychol.*, 1937, **3**, 115–143.

9. RAPAPORT, D., GILL, M. M., & SCHAFER, R. *Diagnostic psychological testing.* Chicago: Year Book Pubs., 1946, Vol. II, 395–459.

10. ROSENZWEIG, S. Apperceptive norms for the Thematic Apperception Test. *J. Pers.*, 1949, **17**, 475–503.

11. ROTTER, J. B. Thematic Apperception Test: suggestions for administration and interpretation. *J. Pers.*, 1946, **15**, 70–92.

12. STEIN, M. I. *The Thematic Apperception Test: an introductory manual.* Cambridge: Addison-Wesley Press, 1948.

13. TOMKINS, S. S. *The Thematic Apperception Test.* New York: Grune & Stratton, 1947.

14. VALENTINE, M., & ROBIN, A. A. Aspects of Thematic Apperception Testing: Depression. *J. ment. Sci.*, 1950, **96**, 435–447.

15. WYATT, F. The scoring and analysis of the Thematic Apperception Test. *J. Psychol.*, 1947, **24**, 319–330.

Received January 18, 1951.

PARENTAL "IDENTIFICATION" AND "TENDENCY TOWARD DISORDERS" AS MEASURED BY THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY[1]

ANDREW L. SOPCHAK
University of Maine

THE purpose of this study was to determine the relation between the tendency towards identification with parents and tendencies toward specific types of abnormalities as measured by the Minnesota Multiphasic Personality Inventory (MMPI).

The term, *identification*, like most terms borrowed from psychoanalysis has been used to designate a variety of meanings. Symonds (5), after surveying the major psychoanalytic works, found that the term was used with three distinct meanings. This study, however, was concerned with Symond's first definition which is as follows: "The most common use of the term is the modeling of oneself in thought, feeling or action after another person." This is almost synonymous with imitation. The sense of "modeling oneself after another person" is of considerable significance for psychological theories concerning the origin of normal personalities in general, and also of specific abnormalities in particular.

Two hypotheses have been stated regarding the relation between identification with parents and the adjustment of the individual. According to one psychoanalytic hypothesis, identification with parents leads to normal character development and adjustment. This point of view would lead one to expect abnormality to develop where there is lack of identification. Another psychoanalytic hypothesis states that identification with maladjusted parents leads to and is a cause of abnormal trends.

This study constituted a partial test of both of these hypotheses. It is, of course, possible that both hypotheses are correct. Identification with one's parents might cause normality in one individual and abnormality in another. On the other hand, identification with one's parents might cause tendencies towards some abnormalities whereas lack of identification might cause tendencies toward others. The exact relationship may vary also with the sex of the individual who is identifying. In brief, if the correlations between identification with parents and the trends toward abnormalities (which are quite numerous) were known, considerable light would be thrown on the validity of certain hypotheses. This study consists of an attempt to determine several such correlations.

METHOD

The Measures

The measures of a trend toward a specific type of abnormality used in this investigation were scores on the MMPI. The higher the score for a particular type of abnormality, the greater it was assumed was the trend toward this type of abnormality.

This assumption is justified on the basis of the data supplied by the authors in their manual for the MMPI. They state that most abnormal subjects score above 70 on one or more of the scales and that the majority of clearly abnormal persons score above 70 on two or more scales. Furthermore, they state that occasionally, when most of the profile shows T-scores of 50 or below, with a single point reaching up above 60, it is safe to interpret this point as if it had reached above 70. These T-scores are standard scores in which the mean of the normative group is assigned a value of 50, and a standard deviation of 10. When only one end of a scale has been identified as abnormal in a clinically recognized sense, the scale is always oriented so that the larger T-scores, i.e., those above 50, represent the abnormal direction.

The normative data used in this study

[1] The writer is grateful to Dr. Luther L. Mays, formerly of the Psychology Department, Syracuse University, at present Chief Clinical Psychologist, Veterans Administration Hospital, Washington, D. C., for his advice and guidance in the study and thesis on which this article is based.

were derived from a sample of about 700 individuals representing a cross section of the Minnesota population as obtained from visitors to the University Hospitals. The sampling was fairly adequate for the ages 16–55 and for both sexes. In addition to these data on normal individuals, data were available for 250 precollege and college students. The scales were developed by contrasting the normal groups with carefully studied clinical cases of which over 800 were available from the neuropsychiatric division of the University Hospitals. The chief criterion of excellence has been the valid prediction of clinical cases against the neuropsychiatric staff diagnosis, rather than statistical measures of reliability and validity. Nevertheless, the reliability coefficients on the scales range between .71 and .83. These correlations were obtained by testing and retesting a group of 40 normals at intervals of a week or less up to several years. A high score on a scale has been found to predict positively the corresponding final clinical diagnosis in more than 60 per cent of new psychiatric admissions. This percentage was derived from differentiation among clinic cases, which is considerably more difficult than differentiation of normal groups.

The measure of *identification* used in this study was the similarity among (a) the answers which the subject gave in taking the MMPI in the usual way (condition A), (b) the answers he gave when taking it with instructions to answer the questions as he thought his father would (condition B) or as his mother would (condition C), and (c) when the subject was given instructions to answer the inventory, substituting "most people" for the other conditions A, B, and C.

A count was made of the number of items in the test which the individual answered in the same manner for himself (condition A) and for his father (condition B). If, for example, he gave exactly the same answer to 325 questions when taking the test for himself as when taking it for his father, his identification score would be 325. This score was assumed to represent the amount of identification which he showed for his father.

It should be pointed out that the identification score obtained by this method represents identification with the subject's "image" of the person with whom he is identifying himself. A little thought will show, however, that this is the only type of identification which can exist. One can identify himself with another only by identifying himself with his image of the other person, because the subject can never know what the other person actually is. It is doubtful whether identification scores obtained by having the inventory answered by the parents themselves and comparing their answers with the answers given by the subjects would yield a more valid test of the relationships studied in this investigation, than would the technique actually used. However, this remains to be seen from the study now in progress.

Subjects

The subjects (*Ss*) participating in this study were 108 students, 78 men and 30 women, who were taking their second semester of elementary psychology at the University of Maine. The data for each of these 108 *Ss* were complete.

Procedure

The topic of personality had been introduced to the *Ss* during their regular classes in psychology. Although such tests as the Bernreuter Inventory had been discussed with them, the MMPI had not.

The group form of the MMPI was administered to the *Ss* in regular laboratory meetings of the psychology class. The successive administrations (conditions A, B, C, and D previously described) were spaced at one week intervals.

In answering the questions the *Ss* were given instructions to answer the questions in terms of how they thought the persons cited in the conditions, ie. "father," "mother," and "most people" would answer them if they were answering the questions for themselves. It was not assumed that the *Ss* could be sure to know how the father, mother, or most people would answer. Therefore, the *Ss* were instructed to guess if necessary.

In order to control, as far as possible, any effect that might have resulted from the sequence in which the *Ss* served under the various conditions, they were divided into four groups which went through the four different conditions in four different sequences as follows:

Experimental Period	Group			
	1	2	3	4
First	A	B	C	D
Second	B	D	A	C
Third	C	A	D	B
Fourth	D	C	B	A

The collected data were scored in such a way as to give 15 scores for each *S*.

a. The inventory was first scored in the normal manner for condition A. This gave 12 scores

for each *S*. Nine of these were the scores for the several categories of abnormality measured by the inventory, the other three were the validity scores "?," "L," "F."

b. Three identification scores were determined for each *S*, namely, identification with father (conditions A and B), identification with mother (conditions A and C), and identification with "most people" (conditions A and D).

This third identification score does not have any direct bearing on the main purpose of the study, namely, the determination of the relation between identification with parents and abnormality. It was secured to serve as a control. It was assumed that the correlations between abnormality scores and identification with "most people" would be lower than the correlations between abnormality scores and identification with father and mother.

It might be assumed that a person going through the process of filling out the test as he thinks some other person would fill it out would tend to give the same answers for the other person as he would for himself owing to his projecting his own habits of thought and feelings into those of others. However, there is no way of being certain whether his answers are projected or whether he is identifying himself with another in answering the test questions.

The eventual verification of the assumption that similarity of answers measures identification rather than projection must await the accumulation of supporting evidence from future studies. But some evidence on the problem can be secured by comparing the relative degree of similarity of answers for the self and parents with that for the self and "most people." Both clinical observation of maladjusted persons and general observation of normal ones would lead one to expect that answering an item in the same way for himself and for his parents results relatively more from identification with the parent than from projection, whereas with reference to answers for "most people," projection should be more operative than identification. Assuming that *identification with parents* is usually stronger than *projection toward "most people,"* higher identification scores should be obtained for the answers for parents than for those for most people. In order to provide a partial check on the validity of this proposition, an identification score with "most people" was obtained.

c. Correlation coefficients were calculated between each of the nine scores of abnormality and the three identification scores thus giving a total of 54 correlation coefficients, 27 for the men and 27 for the women. These coefficients constitute the data bearing upon the purpose of the study.

RESULTS

It can be seen from Table 1 and Figure 1 that the means of the group of *S*s used all fell within the "normal" range of all scales. (scores falling between T-scores of 30 and

TABLE 1

MEANS AND STANDARD DEVIATIONS IN RAW SCORE UNITS AND T-SCORE UNITS OF 78 MALE AND 30 FEMALE SUBJECTS ON THE VARIOUS ABNORMALITY SCALES OF THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY

SCALE	*Hs*	*D*	*Hy*	*Pd*	*Mf*	*Pa*	*Pt*	*Sc*	*Ma*
Men									
Mean raw score	11.7	20.4	19.1	22.0	24.5	8.0	26.9	25.2	20.5
SD	3.4	5.1	4.5	4.3	5.2	2.9	5.9	5.3	3.5
Mean T-score	52	58	55	57	50	50	58	55	60
SD	8.7	12.2	8.1	10.1	10.1	8.4	12.1	10.2	8.7
Women									
Mean raw score	11.7	19.5	19.7	19.8	37.3	8.4	28.2	25.0	20.6
SD	2.9	5.1	4.2	3.9	5.2	3.1	4.7	3.6	4.1
Mean T-score	48	51	52	53	49	50	55	54	60
SD	5.9	9.9	7.4	9.0	10.6	9.0	7.7	5.1	11.9

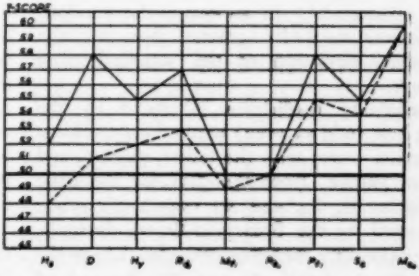

FIG. 1. MEANS IN T-SCORE UNITS OF 78 MALE AND 30 FEMALE SUBJECTS ON THE VARIOUS ABNORMALITY SCALES OF THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY

Males ————

Females - - - - - - - -

70 are within the normal range). Thus, this range of normality indicates that the scores on the abnormality scales are valid according to the criteria established by the authors of the test. The mean scores for the women are somewhat nearer the mean score of 50 found in the standardizing group used by the authors of the test than those of the men. Nevertheless, the means for both groups were clearly normal. For both men and women the standard deviations are in the majority of cases smaller than those of the standardizing group used by the authors of the test (their group included both normals and abnormals). In no case is the standard deviation markedly greater than 10, the standard deviation of the standardization group.

Reliability coefficients calculated by the split half (odd vs. even) method for the identification scores with the father, mother, and "most people" from the data for the male *Ss* were .81, .83, and .97 respectively when corrected by the Spearman-Brown formula. Corresponding coefficients calculated from the data for the female *Ss* were .89, .83, and .96 respectively.

Table 2 indicates that men who possess tendencies toward abnormality show, in general, a lack of identification with their fathers. This relationship is not statistically significant for all types of abnormality, but is very significant ($p<.01$) for tendencies toward femininity, paranoia, psychasthenia, schizophrenia, and hypomania. Men who possess tendencies toward abnormality also show a lack of identification with their mothers though this failure to identify with the mother is not as marked as failure to identify with the father. The relationship is very significant ($p<.01$) for tendencies toward femininity, psychasthenia, schizophrenia, and hypomania, and it is significant at the .05 level for depression. Tendencies toward abnormality in men are also accompanied by a lack of identification with "most people," but here the relationship is less marked than for lack of identification with either father or mother. The correlations are very significant ($p<.01$) only for tendencies toward femininity, schizophrenia, and hypomania and are significant at the .05 level for only depression and paranoia.

Considering Table 2 as a whole, it is obvious that the correlation between a lack of identification and tendencies toward abnormality decreases as one considers identification with father, mother and "most people" in sequence. Although the difference between identification with father and with

TABLE 2

PEARSON COEFFICIENTS OF CORRELATION BETWEEN THE VARIOUS SCALES OF THE MMPI AND THE IDENTIFICATION SCORES
(N=78 male *Ss*)

Inventory Character	FATHER			MOTHER			MOST PEOPLE		
	r	σ_r	*t*-ratio	*r*	σ_r	*t*-ratio	*r*	σ_r	*t*-ratio
Hs	−.12	.112	1.1	−.04	.114	0.4	−.19	.110	1.7
D	−.20	.109	1.8	−.25*	.107	2.2	−.27*	.106	2.4
Hy	−.07	.113	0.6	−.09	.113	0.8	−.08	.113	0.7
Pd	−.09	.114	0.8	−.13	.112	1.1	−.16	.111	1.4
Mf	−.46**	.090	4.0	−.41**	.095	3.6	−.32**	.102	2.8
Pa	−.33**	.101	2.9	−.20	.109	1.7	−.27*	.106	2.4
Pt	−.44**	.092	3.9	−.34**	.101	3.0	−.21	.109	1.8
Sc	−.37**	.098	3.2	−.34**	.101	3.0	−.36**	.099	3.2
Ma	−.39**	.097	3.4	−.40**	.096	3.5	−.34**	.101	3.0
Identification Score									
Mean	432.1			419.2			433.3		
SD	54.6			48.3			43.3		

* Significant at .05 level ($r \geqq .2234$).
** Significant at .01 level ($r \geqq .2906$).

mother is not statistically significant, there is a higher identification with the father for this group as a whole. At the .01 level five of the correlations with abnormality scores are significant for identification with father, four with mother and three with "most people." The exact size of the coefficients decreases in the same direction.

Study of the relative size of the correlations in this table gives information which is of significance to a study by Gough (4) who analyzed the scores of abnormal patients on the various scales of the MMPI and found that the scores of neurotic and psychotic patients tended to form distinct and characteristic patterns. He describes the psychoneurotic curve having the "neurotic triad" comprised of hypochondriasis, depression, and hysteria. He found the scores for neurotic subjects to be higher on this triad than on the "psychotic triad" consisting of paranoia, psychasthenia, and schizophrenia. The score pattern for the psychotics were just reversed, being high on the "psychotic" and low on the "neurotic triad."

In the present study the correlations are negligible on the "neurotic triad" whereas the "psychotic triad" correlations show the existence of a substantial relationship. This relationship is especially marked in the case of failure to identify with the father. This would imply that the association between abnormal tendencies and lack of identification is most marked in the cases of the psychoses as contrasted with the neuroses, and that this association with psychoses is most marked for failure to identify with father and less for identification with mother and "most people." The correlations for hypomania are more similar to those in the "psychotic triad" than to those in the "neurotic triad."

TABLE 3

PEARSON COEFFICIENTS OF CORRELATION BETWEEN THE VARIOUS SCALES OF THE MMPI AND THE IDENTIFICATION SCORES
(N=30 female *Ss*)

Inventory Character	Father			Mother			Most People		
	r	σ_r	*t*-ratio	*r*	σ_r	*t*-ratio	*r*	σ_r	*t*-ratio
Hs	.02	.185	.11	.27*	.173	1.45	.03	.186	.16
D	−.25*	.174	1.34	.03	.186	.16	.05	.186	.27
Hy	.21	.178	1.13	.12	.183	.65	.16	.181	.86
Pd	−.03	.186	.16	.05	.186	.27	−.07	.185	.09
Mf	−.47**	.145	2.51	.11	.183	.59	.15	.181	.81
Pa	−.10	.184	.54	.17	.181	.91	.00	.000	.00
Pt	−.41**	.155	2.20	.12	.183	.65	.03	.186	.16
Sc	−.42**	.153	2.30	.13	.183	.70	.03	.186	.16
Ma	−.15	.181	.81	.30**	.169	1.61	.01	.186	.05
Identification Score									
Mean	423.5			430.1			420.9		
SD	36.5			29.9			96.9		

* Significant at .05 level ($r \geqq .2234$).
** Significant at .01 level ($r \geqq .2906$).

If consideration is now given to the data for the female subjects, it can be seen in Table 3 that women who possess tendencies toward abnormality show a lack of identification with their fathers which is significant at the .01 level in the case of the masculinity scale only. The same relationship is significant at the .05 level in psychasthenia and schizophrenia. There are positive but nonsignificant correlations between identification with mother and all types of abnormality trends. What little indication of identification with most people does exist points toward positive identification for all types of abnormality excepting that for psychopathic deviance.

Considering Table 3 as a whole it is obvious that the relation between identification

and abnormality is quite different in men from that in women in so far as the data of this study are valid for people in general. Women with tendencies towards psychasthenia and schizophrenia lack identification with the father as was found to be true for men, but the correlations are significant at only the .05 level, whereas for men they are significant at the .01 level. This finding is difficult to interpret since it indicates that "masculine" women identify less with their fathers than do feminine women.

Speculation as to the cause of the lack of identification with the father exhibited by those with abnormal tendencies might give rise to innumerable interpretations. One of the most plausible is that the individual who feels hostile towards his father fails to identify with him as a result of this feeling of hostility. Such an interpretation might explain the seeming incongruity that women tending towards masculinity fail to identify with their fathers. It is possible that masculine women hate and compete with their fathers, and that even though they may possess masculine traits in common with them, by and large they possess traits different from their fathers and consequently show a lack of identification with them.

For other abnormal tendencies, however, there is no significant indication of lack of identification being associated with the abnormality. As far as identification with mother is concerned, women who show tendencies toward abnormality may identify with the mother as contrasted with the situation existing in men where abnormality was accompanied by lack of identification. This positive identification with the mother in women is not statistically significant, however.

There is apparently no significant relation between identification with most people and abnormality for women whereas men with abnormal trends show some lack of identification.

There is one noticeable similarity between the data for men and women, namely, that the association between failure to identify with the father and tendency toward the "psychotic triad" of Gough tends to hold true for women as well as for men. However, with men this relationship is most obvious in the case of identification with father as contrasted with that of the mother and "most people."

It should be pointed out that the data for men and women both show noticeably smaller correlations for identification with "most people" than for identification with father or mother. This is in agreement with an assumption made previously that this relationship would exist if the technique used measured identification more than projection. The data for women clearly justify the assumption that the technique measures identification more than do those for the men. The presence of significant correlations for identification with "most people" in the data for the men does not, however, necessarily indicate that the technique does not measure identification. Even though their correlations are less for "most people" than for mother or father, it is just as plausible to assume that men with abnormal trends fail to identify with people in general as it is to assume that the correlations are due to projection rather than identification. The final outcome or answer as to the validity of the technique will, of course, have to await further research. Additional evidence favoring the validity of the technique is that for both men and women the mean identification scores indicate greater identification with the parent of the same sex than with the parent of the opposite sex.

SUMMARY AND CONCLUSIONS

Determining the relation between the tendency towards identification with parents and tendencies toward the specific types of abnormalities as measured by the MMPI, through the technique of repeated administration of the MMPI under different instructions, has resulted in a number of conclusions:

1. As a group, the normal men tend to identify with their fathers more than with their mothers.

2. The men in the group with tendencies toward abnormality show a greater lack of identification with their fathers than with their mothers.

3. Those men in the group with tendencies toward abnormality also fail to identify with their mothers.

4. Tendency toward abnormality in men is accompanied by lack of identification with people in general.

5. For those men who fail to identify with the father, tendencies toward the abnormalities of the "psychotic triad" will be stronger than the tendencies toward the "neurotic triad."

6. Women with tendencies toward abnormality show a lack of identification with their fathers but this lack of identification is not as strong as in the case of men.

7. "Masculine" women identify with their fathers less than feminine women.

8. For women, positive identification with the mother apparently is a correlate of some types of abnormality. This suggests that identification with the father is more important in producing normal adjustment than is identification with the mother.

9. For women no significant relation between identification with "most people" and abnormality exists.

10. The association between failure to identify with the father and tendencies toward the "psychotic triad" holds for the women.

11. For both men and women, the mean identification scores indicate greater identification with the parent of the same sex than with the parent of the opposite sex.

12. For both men and women failure to identify with the father is more closely associated with trends toward abnormality than is failure to identify with the mother.

13. The data for both men and women as a whole seem to favor the hypothesis that identification with parents is associated with normality and good adjustment rather than the opposing hypothesis that identification is associated with abnormal trends. This statement holds true for normal persons as a group in so far as the subjects used in this study represent normals in general. It is, of course, possible that the hypothesis of a positive association between identification with parents and abnormal trends holds for given individuals. The data obviously suggest that the last mentioned hypothesis is more likely to be true for women identifying with their mothers than for men identifying with either mother or father, or for women identifying with their fathers.

14. This study, as a whole, indicates that tendencies toward abnormality as measured by the MMPI are in general associated with failure to identify with the parents, especially in the case of men. This general tendency also holds somewhat for women in so far as identification with the father is concerned.

REFERENCES

1. CAMERON, N. *The psychology of behavior disorders.* Boston: Houghton Mifflin, 1947.
2. FREUD, S. *New introductory lectures on psychoanalysis.* New York: Liveright, 1935.
3. GOUGH, H. Simulated patterns on the Minnesota Multiphasic Personality Inventory. *J. abnorm. soc. Psychol.*, 1947, **42**, 215–225.
4. GOUGH, H. Diagnostic patterns on the Minnesota Multiphasic Personality Inventory. *J. clin. Psychol.*, 1946, **2**, 23–27.
5. SYMONDS, P. *The dynamics of human adjustment.* New York: D. Appleton-Century, 1946.

Received February 5, 1951.

A STUDY OF ANXIETY AND LEARNING

GEORGE MANDLER AND SEYMOUR B. SARASON
Yale University

THE present study is an attempt to investigate the role of drive states in a testing situation; more specifically, the extent to which anxiety responses are evoked in the situation and the relation of such responses to performance and learning. In addition to studying the role of anxiety, consideration is given to the effects of subjective success and failure on the anxiety state and test performance.

Comparatively little attention has been paid in the literature to the influence of various drive states on the performance of typical intelligence test items. Thus, Mursell (8) mentions motivation only in comparing educational with mental tests, and Terman and Merrill (10) refer briefly to "hunger restlessness, desire to please . . ., timidity and a hundred other motives. . . ."

Diethelm and Jones (3) found that the presence of clinical anxiety significantly decreased scores on the Kohs Block Design Test for most of their adult subjects, and that maze learning was reliably slower under anxiety conditions. They conclude that the "possible influence of anxiety . . . in psychological examinations must be considered more seriously." Malmo and Amsel (6) investigated the effect of neurotic anxiety upon rote learning and demonstrated the role of anxiety-produced interference upon learning. They concluded that the "forgetfulness of the anxious patient is due to anxiety-produced interference between the relevant responses and the irrelevant responses generated out of the patient's anxiety state."

In the present study, in contrast to previous ones (1, 6), the measure of anxiety was based on a questionnaire which was specifically concerned with the subjects' attitudes and experiences in a testing situation.

THEORETICAL CONSIDERATIONS AND HYPOTHESES

In a previous paper Sarason (9) has pointed out that in both theoretical and practical approaches to the testing situation insufficient attention has been given to the nature, role, and strength of internal or covert drive states. In the present investigation certain aspects of the testing situation are emphasized: the drive states and the resultant responses. The following assumptions refer to the drives evoked and types of responses elicited in the test situation. These drives will be considered as falling into two major categories:

Assumption 1. Learned drives which are a function of the nature of the task, test materials, and instructions. These include the need to achieve and to finish the task; in short, drives which evoke responses relative to satisfying the requirement set by the task or the experimenter. We shall refer to these as task drives (S_T). It is assumed that these drives are reduced by task responses (R_T), which are responses or response sequences which lead to completion of the task.

Assumption 2. A learned anxiety drive which is a function of anxiety reactions previously learned as responses to stimuli present in the testing situation. Anxiety is here considered as a response-produced strong stimulus with the functional characteristics of drives as discussed by Miller and Dollard (7). Anxiety reactions are generalized from previous experiences to testing situations. The anxiety drive (S_A) primarily elicits responses that tend to reduce the drive. These responses are considered to be of two general types:

(a) Anxiety responses which are not specifically connected with the nature of the task or materials. These responses (designated as R_A) may be manifested as feelings of inadequacy, helplessness, heightened somatic reaction, anticipations of punishment or loss of status and esteem, and implicit attempts at leaving the test situation. It might be said that these responses are self rather than task centered.

(b) Anxiety responses which are directly related to the completion of the task and which reduce anxiety by leading to completion of the task. These responses (designated R_{AT}) are functionally equivalent to R_T responses. More specifically, intervening

responses r_{at} and r_t lead to the final response R_T.

Assumption 3. Responses of the type R_T and R_{AT} lead, in point of time, to task completion, while R_A type of responses interferes with task completion. It is assumed that intervening r_a responses are incompatible with r_t and r_{at} responses.

Assumption 4. R_{AT} responses, which are considered to be specific to the task, are not available in the response repertory of an individual, but are evoked and learned in the course of task performance. R_A responses, which are not considered to be specific to the task, are available in the response repertory and are, by the process of generalization, readily evoked. Thus, individuals with a high anxiety drive and a large number of R_A anxiety responses in their response repertory will tend to make more R_A responses initially than individuals with a low anxiety drive and, conversely, in relation to the total number of responses available, individuals with a low anxiety drive will tend to make more task relevant anxiety responses (R_{AT}) than those with a high anxiety drive.

Figure 1 is a diagrammatic presentation of the drives and responses evoked in the testing situation.

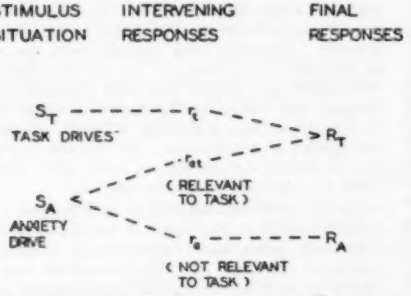

Fig. 1. Diagrammatic Presentation of Drives and Responses Evoked in the Testing Situation

These assumptions lead to the following hypotheses:

Hypothesis 1. When an anxiety drive (S_A) evokes intervening responses predominantly of the nature r_{at}, learning will, because of summation of drives (S_T and S_A), proceed at a faster rate than under the condition when incompatible r_a and r_{at} responses are evoked.

Hypothesis 2. When a task is completed in each trial and when improvement in succeeding trials is an easily observable cue, more responses of the order R_{AT} will be reinforced than responses of the order R_A. Such final task completion responses (R_{AT}) will be anxiety reducing and will become progressively stronger and dominant.

Hypothesis 3. When improvement in a task from trial to trial is not readily apparent to the subject, responses of the order R_{AT} will not be substantially reinforced and more random responses of the order R_A will be evoked, especially if the anxiety drive is strong to begin with.

Hypothesis 4. When the dominant response tendency has been of the order R_A (responses not relevant to the task), then any further reference to the test situation (such as success or failure) will evoke predominantly R_A responses.

Hypothesis 5. When there is no initial tendency to make R_A responses and when R_{AT} responses have been predominantly evoked in a task, then presentation of a new but similar task preceded by a reference to the test situation will elicit further R_{AT} responses.

Method

Section 1

One hundred and fifty-four students in an introductory psychology course in Yale College were given an anxiety questionnaire.[1] The questionnaire was presented to the students as a "questionnaire on attitudes toward test situations." It consisted of 67 questions and was divided into four sections: (a) individual intelligence tests, (b) group intelligence tests, (c) course examinations, and (d) general questions.

Each section contained questions ("anxiety questions") dealing with the student's subjective experiences in the testing situation such as uneasiness, accelerated heartbeat, perspiration, emotional interference, and "worry" before and during a testing session. The questionnaire also contained questions relating to attitudes (likes and dislikes) toward tests. The *Ss* were requested to mark their answers for each question anywhere along a 15-centimeter, graphic scale with endpoints and midpoint indicated. Only nonveteran sophomores and juniors who completed the questionnaire were included in this study. The resulting 101 questionnaires were then scored by assigning a score, expressed in millimeters, to each question. The resulting 101 scores for each question were then tabulated and a subject was given a score of 0 if

[1] The authors wish to thank Dr. Claude E. Buxton and the students of his introductory psychology course for their wholehearted cooperation.

he fell below and a score of 1 if he fell above the group median. A total score for each subject based on the 0 and 1 scores was then determined for 42 anxiety questions and 25 attitude questions. For the purpose of this study only the anxiety scores were used. The distribution for each of the 42 questions was then compared with the distribution based on all 42 questions. For this purpose the subjects were divided into two groups at the total score median, and the consistency with which each of the 42 questions divided the subjects into these two groups was determined by chi-square. All questions which divided subjects on this basis at better than chance expectancy (at the .03 level of confidence) were retained. Five questions did not fulfill this condition and were discarded. From the remaining 37 questions each subject obtained a revised score theoretically ranging from 0 to 37. The split-half reliability (odd vs. even questions) of the anxiety questionnaire was .91 (Spearman-Brown prophecy estimate).

The final anxiety questionnaire scores yielded a distribution of the 101 subjects along a presumed anxiety continuum. Forty-two subjects with the 21 lowest and 21 highest scores were then selected for the actual study, i.e., one high-anxiety group (designated HA) and one low-anxiety group (LA). Each of these two groups was divided by random procedure into 3 subgroups (Success, Failure, and Neutral). Approximately 3½ months after the administration of the questionnaire these students were asked to participate in a psychological experiment concerning intelligence tests. Apparently none of the subjects was aware of any connection between the questionnaire and the testing session.

Section II

At the beginning of the testing session the subject was told that he was going to be given a number of intelligence tests as part of a study relating to the aptitude tests given to every Yale freshman. He was also told that his scores in the session would be compared with his aptitude test scores to make possible a better interpretation of the latter. He was presented with the Kohs Block Design Test and asked to do sample design #2 of the Wechsler-Bellevue modification. He was then presented with design #13 from the Kohs Block Design Test (5) and the total time (in seconds) needed for completion was recorded. The subject was then presented with the Digit Symbol subtest from the Wechsler-Bellevue Intelligence Test (11), doing the practice part first. A time limit of one minute was used for the main part, scoring number right minus number wrong. The subject was then again presented with Kohs Design #13 and, after that, again with the Digit Symbol test. This was continued until each *S* had completed six trials of each of the tests. Six trials were chosen in order to obtain both a representative performance curve and to minimize boredom that could appear with a large number of trials. After the completion of these trials each subject was given the Wisconsin Card Sorting Test[2] (2, 4) and informed at its conclusion that this was not part of the intelligence testing but an additional, purely exploratory, study.

Section III

At this point the experimenter secretly opened an envelope, previously prepared for each subject, informing the experimenter as to which of the three experimental subgroups (success, failure, or neutral) the subject belonged. The procedure was then as follows:

Success group. The subject was informally told that he had done extremely well on the Kohs Block Design and Digit Symbol tests. The experimenter looked through several sheets of paper (apparently score and norm data), leafed through Wechsler's *The Measurement of Adult Intelligence,* and finally informed the subject that he had scored at approximately the 93rd to 96th percentile for a comparable college population. This was concluded by the experimenter's saying: "Let's see whether you can do as well in the second part."

Failure group. The procedure was the same as for the success group except that the subject was told that he had done much worse than would be expected from his aptitude scores, and that his performance fell below the 10th percentile. The experimenter then said: "Let's see whether you can do better in the second part."

Neutral group. The experimenter only said: "Let's go on to the second part."

Section IV

Each subject was then presented with six trials of the Kohs Block Design #16 and a specially prepared Digit Symbol test of the same form as the Wechsler-Bellevue subtest but with different symbols. Procedure and scoring were the same as described in Section II. At the conclusion of the session each subject was fully informed of the purpose and design of the experiment and asked to fill out a short self-rating schedule on anxiety and quality of performance. He was also requested not to discuss the experiment with other students until the end of the testing sessions.

Of the initial 42 subjects, 9 (5 HA and 4 LA subjects) failed to participate in the testing sessions. The final design and distribution of subjects are presented in Table 1. The procedure was designed in such a way that the experimenter could not know to which of the two anxiety groups any subject belonged.

RESULTS AND DISCUSSION

In an attempt to establish face validity of the questionnaire, the experimenter (without knowledge of the subject's questionnaire standing) rated each subject's behavior on a five-point scale according to five criteria of overt anxiety manifestation (perspiration, ex-

[2] The analysis of the subjects' performance on this test will not be presented in this study.

TABLE 1

Experimental Design and Distribution of Subjects (N)

	High-Anxiety Group (N=18)			Low-Anxiety Group (N=15)		
Part I	6 trials Kohs Block Design Test (Design No. 13) 6 trials Digit Symbol Test (Wechsler-Bellevue subtest)					
Part II	Wisconsin Card Sorting Test					
Part III **	HA * Success (N=5)	HA Failure (N=6)	HA Neutral (N=7)	LA Success (N=5)	LA Failure (N=5)	LA Neutral (N=5)
	6 trials Kohs Block Design Test (Design No. 16) 6 trials Digit Symbol Test (new form)					

* HA=High anxiety group, LA=Low anxiety group.
** The *Ss* in each of the two anxiety groups were assigned by random procedure to the 3 subgroups (success, failure, neutral). For the purpose of statistical analysis only five *Ss* in each subgroup were used.

cessive movement, inappropriate laughter and exclamations, questioning of instructions, hand movement). These ratings were dichotomized into a high anxiety and a low anxiety group and compared with questionnaire ratings. The agreement was significant at better than the .001 level of confidence with a point correlation (phi) of .59.

In line with the analysis presented in the introduction, the high anxiety group can be characterized as individuals with a high anxiety drive in testing situations and these subjects would be expected to react to the testing situation with a large variety of intervening responses of the order r_a (responses not relevant to the task). For the low anxiety group the anxiety drive is much weaker.

The following predictions arise out of the assumptions and hypotheses stated previously:

Prediction 1 (Hypotheses 1 and 2). The high anxiety group (HA) is expected to have higher mean time scores and greater variability than the low anxiety group (LA) on the Kohs Block Design Test because of the evocation of more responses of the order R_A. Table 2A and Fig. 2 and 3 show the mean scores and variability for the two groups in the six trials of Part I. The mean score differences all lie in the predicted direction, *p*-values for the first five trials ranging from .04 to .14. The variability of the HA group is significantly larger at better than the .001 level of confidence for trials 2 to 5. Trials 1 and 6 will be discussed separately below.

TABLE 2A

Mean Time Scores and Standard Deviations for Both Anxiety Groups for Six Trials of First Kohs Design

Trial	High Anxiety Group		Low Anxiety Group					
	Mean	SD	Mean	SD	*t* (df=31)	*p* *	*F*	*p* *
1	199.4	114	158.2	107	1.07	.14	1.15	>.05
2	111.1	78	82.6	25	1.36	.08	9.84	<.001
3	76.1	36	65.1	12	1.13	.13	8.83	<.001
4	76.2	35	58.9	12	1.86	.04	8.06	<.001
5	74.4(60.5)**	63(30)**	55.4	11	1.17	.12	32.88	<.001
6	58.9	20	54.3	12	0.78	.22	2.44	.05

* One tail of the distribution of *t* used.
** Figures in parentheses exclude one extreme case, see also Fig. 2 and 3. The respective *t* and *F* values are:
t=0.94 p=.17
F=7.49 p=.001

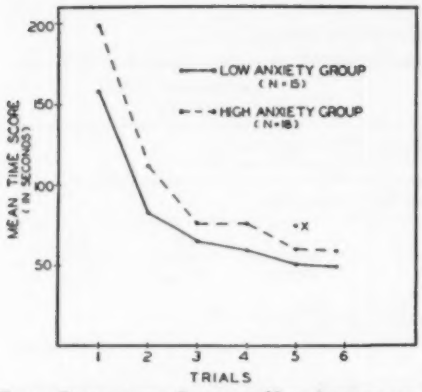

FIG. 2. PERFORMANCE CURVE FOR HIGH ANXIETY AND LOW ANXIETY GROUPS FOR SIX TRIALS OF FIRST KOHS BLOCK DESIGN. (X=MEAN SCORE FOR HIGH ANXIETY GROUP INCLUDING ONE EXTREME CASE.)

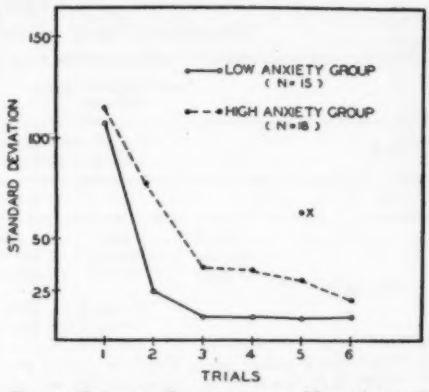

FIG. 3. STANDARD DEVIATIONS FOR HIGH ANXIETY AND LOW ANXIETY GROUPS FOR SIX TRIALS OF FIRST KOHS BLOCK DESIGN. (X=STANDARD DEVIATION FOR HIGH ANXIETY GROUP INCLUDING ONE EXTREME CASE.)

Prediction 2 (Hypotheses 1 and 3). The mean scores on the Digit Symbol test (where high scores indicate better performance) are expected to be lower for the HA group in line with Prediction 1. However, variability is expected to rise for the HA group in line with Hypothesis 3 since the Digit Symbol test was never completed by any subject and improvement in this test was slight for successive trials. The variability of the HA groups should increase because of the absence of anxiety reduction resulting from completion of the task at hand. Table 3 and Fig. 4 and 5 show the scores for the Digit Symbol test. The mean scores for the first five trials result in differences which are significant at the .05 to .17 level of confidence. Figure 5 shows the predicted rise in variability for the HA group accompanied by an apparent drop in variability for the LA group. None of the differences in variability between the two groups was significant. In order to test the difference between the two anxiety groups in variability change from trial 1 to trial 6, the following statistical procedure was followed. In the HA group the 18 *Ss* were assigned randomly to one of six groups. For each of these groups of three subjects an estimate of variability was obtained for trials 1 and 6. The variability on trial 6 was then subtracted from that on trial 1, resulting in six values. The same procedure was employed

TABLE 2B

MEAN TIME SCORES FOR THE EXPERIMENTAL SUBGROUPS ON THE FIRST TRIAL OF SECOND KOHS DESIGN **

	HIGH-ANXIETY GROUP		LOW-ANXIETY GROUP		t	p *
	MEAN	ADJUSTED MEAN	MEAN	ADJUSTED MEAN		
Neutral	104.8	126.0	143.4	165.6	1.46	.08
Success	197.6	163.6	106.0	138.0	0.94	.17
Failure	204.8	162.2	146.2	108.6	1.98	.03

* See footnote to Table 2A.

** Tables 2B and 2C are based on an analysis of covariance of the time scores on trial 1 of the second Kohs design for the six experimental subgroups. The *Ss*' scores on trial 1 of the first Kohs design were used as a statistical control for initial differences in Kohs Block Design performance. The resulting t-values have 23 degrees of freedom.

TABLE 2C

Significance of Difference between Mean Time Scores of Experimental Subgroups Within the Two Anxiety Groups **

	High-Anxiety Group		Low-Anxiety Group	
	t	p *	t	p *
Neutral vs. success	1.38	.09	1.02	.16
Neutral vs. failure	1.34	.09	2.10	.02
Success vs. failure	0.05	...	1.01	.16

* See footnote to Table 2A.
** See footnote to Table 2B.

for the LA group, except that here there were five groups with three subjects in each. A t-test of the difference in variability change between the HA and LA groups was significant at the .04 level of confidence. Thus, when compared with the low anxiety group, the high anxiety group showed a significant increase in variability.

Prediction 3 (Hypotheses 1 and 2). In the Kohs Block Design Test, where each trial was necessarily concluded by successful completion and where improvement is easily apparent, R_{AT} responses are strongly reinforced and become increasingly stronger as compared to R_A responses elicited by the anxiety drive. It can be predicted that eventually the high anxiety group will be equal to or better than the low anxiety group because of the summation of drives S_A and S_T. This is partly confirmed by trial 6 of the first part, where the difference between the two groups is least apparent. However, when the Neutral groups (no intervening stimulation between parts I and III) are compared, this predicted difference is striking. The HA group, which has been performing at a lower level than the LA group for the six previous trials, has a mean time score which is better (at the .08 level of confidence) than the LA group on trial 1 of the second design.

Prediction 4 (from Hypothesis 4). For the HA group any further reference to the testing situation in terms of success or failure will elicit internal responses which will increase the strength of the anxiety drive and evoke random responses not relevant to the task (R_A). Thus, the neutral HA group is expected to perform better than either the failure or success group. Their mean time score on the Kohs Block Design Test is lower, but this difference is not highly significant (at the .09 level of confidence). There

TABLE 3

Mean Scores and SD's for Both Anxiety Groups for Six Trials of Digit Symbol Test (Wechsler-Bellevue Subtest) with Time Limit of One Minute

Trial	High Anxiety Group		Low Anxiety Group		t (df=31)	p *
	Mean	SD	Mean	SD		
1	38.5	4.4	40.3	6.6	0.95	.17
2	41.8	5.2	44.3	7.0	1.19	.12
3	44.4	5.3	47.7	6.3	1.64	.05
4	47.9	5.5	49.9	5.9	1.00	.16
5	49.3	5.7	51.5	5.8	1.09	.14
6	51.3	7.2	52.3	5.3	0.45	.33

* One tail of the distribution of t used.
(None of the F-values testing the significance of difference of the variances of the two groups on the first 6 trials was significant (p values $>.05$ throughout). None of the t- or F-values comparing the first trials of the second Digit Symbol test of the 6 subgroups was significant.)

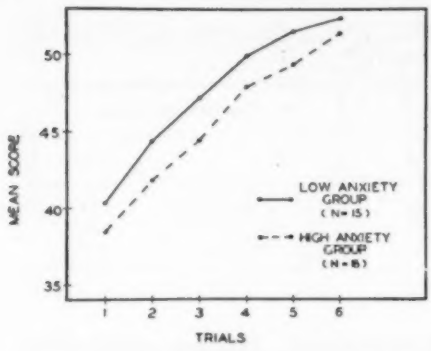

FIG. 4. PERFORMANCE CURVE FOR HIGH ANXIETY AND LOW ANXIETY GROUPS FOR SIX TRIALS OF DIGIT SYMBOL TEST (WECHSLER-BELLEVUE SUBTEST, TIME LIMIT OF ONE MINUTE)

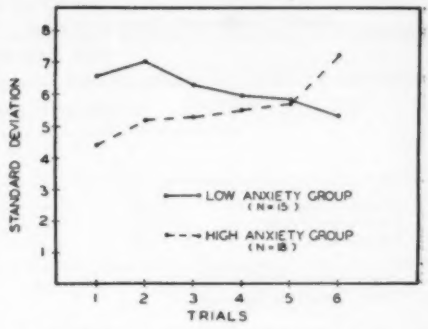

FIG. 5. STANDARD DEVIATIONS FOR HIGH ANXIETY AND LOW ANXIETY GROUPS FOR SIX TRIALS OF DIGIT SYMBOL TEST (WECHSLER-BELLEVUE SUBTEST, TIME LIMIT OF ONE MINUTE)

is no difference, or trend toward a difference, between the success and failure groups.

Prediction 5 (from Hypothesis 5). Any further reference to the testing situation in terms of success and failure is expected to improve the performance of the LA group due to the summation of drives S_A and S_T. Table 2C shows that the difference between the neutral and failure groups is significant (at the .02 level of confidence), whereas the mean time score of the success group lies between the other two groups.

Further Observations

A comparison of the two failure subgroups shows that the low anxiety group is significantly better than the high anxiety group (at the .03 level of confidence). This also lends support to the hypothesis that presumably anxiety-evoking stimulation can have a differential effect (interfering or supporting) on different individuals.

It appears that, in general, the high anxiety group performs better when no further reference is made to their performance, while the low anxiety group profits by such references, apparently in the direction of increased motivation.

The lack of significant differences between the two groups on trial 1 of the first Kohs Design could be attributed to the factor of unfamiliarity which results in large variability for the two groups. However, the null hypothesis was tested that the HA and LA groups are not differentiated by the over-all median when the two groups are thrown together. A chi-square test allowed the rejection of the null hypothesis at the .03 level of confidence.

No data have been given as to the effect of the different types of report on performance on the Digit Symbol test. All the relevant comparisons were not significant, which may be due to a lack of sensitivity and small range of scores on this test.

CONCLUSIONS

The results of the present study suggest that anxiety present in the testing situation is an important variable in test performance. It is questionable whether intelligence test scores adequately describe the underlying abilities of individuals who have high anxiety drive in the testing situation, particularly since the relation of the type of the test to the test performance seems to play an important determining role. Furthermore, anxiety does not necessarily depress scores, but can serve to elicit improvement. Further study of the nature of the various anxiety responses is necessary, particularly the two types of responses (compatible and incompatible with test responses) suggested in this study. In demonstrating the role of anxiety-produced interference, no attempt has been made to study the origin and establishment of the interfering anxiety responses.

Further research into the factors discussed above is necessary and at present is being undertaken by the authors. Among the more important variables being subjected to further scrutiny is the specific effect of instructions, i.e., does the test situation per se produce the differences between the two groups or is this difference a function of specific instructions given by the examiner? Another variable subject to further experiment is the factor of finishing a task and improving on successive trials. Furthermore, the influence of familiarity with the test material and its effect upon differences between the two anxiety groups appears to be an important variable in interpreting test results in the light of the data presented in this study. Another problem presented by our data is the differential effect of success and failure. It is presumed that this effect is to some degree dependent upon *S*'s own evaluation of his performance. Finally, it is suggested that an examination of *S*'s previous experiences and performances in testing situations, as well as specific personality factors and life history data, may be valuable in shedding further light on this problem.

Summary

The present study was designed to investigate the influence of anxiety, as evoked by a testing situation, on the performance of typical intelligence test items. Relevant hypotheses were postulated on the basis of S-R behavior theory.

Two groups of *S*s (a high anxiety group and a low anxiety group), all sophomore and junior college students, were selected on the basis of an anxiety questionnaire and each group was randomly divided into three subgroups (success, failure, and neutral). All subjects were given six trials each on the Kohs Block Design #13 and the Digit Symbol subtest of the Wechsler-Bellevue Intelligence Test. The subjects were then told either that they had done very well, or very badly, or were told to go on to the second part. In the second part all *S*s were give six trials each of the Kohs Block Design #16 and a comparable variation of the Digit Symbol Test.

The following predictions based on the theoretical hypotheses were borne out:

1. The mean time scores on the Kohs Block Design of the low anxiety group were better than those of the high anxiety group for the first five trials.

2. The variability of the high anxiety group was significantly larger than that of the low anxiety group.

3. As the learning process proceeded, the anxiety drive of the high anxiety group tended to improve performance scores.

4. An intervening report (success or failure) elicited improved performance for the low anxiety group but depressed scores for the high anxiety group.

It appears that the optimal conditions for a high anxiety group are those in which no further reference is made to the testing situation, and that the optimal conditions for a low anxiety group are those in which the subjects are given a failure report.

The above processes were discussed in terms of responses relevant to the task and responses not relevant to the task, both types being evoked by an anxiety drive.

Further research problems arising out of the present study are discussed.

References

1. Beier, E. G. The effect of induced anxiety on some aspects of intellectual functioning; a study of the relationship between anxiety and rigidity. *Amer. Psychologist*, 1949, 4, 273–274. (Abstract.)
2. Berg, E. A. A simple objective technique for measuring flexibility in thinking. *J. gen. Psychol.*, 1948, 39, 15–22.
3. Diethelm, O., & Jones, M. R. Influence of anxiety on attention, learning, retention and thinking. *Arch. Neurol. Psychiat.*, Chicago, 1947, 58, 325–336.
4. Grant, D. A., & Berg, E. A. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *J. exp. Psychol.*, 1948, 38, 404–411.
5. Kohs, S. C. *Intelligence measurement.* New York: Macmillan, 1923.
6. Malmo, R. B., & Amsel, A. Anxiety-produced interference in serial rote learning with observations on rote learning after partial frontal lobotomy. *J. exp. Psychol.*, 1948, 38, 440–454.
7. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
8. Mursell, J. L. *Psychological testing.* (2nd Ed.) New York: Longmans, Green, 1949.
9. Sarason, S. B. The test situation and the problem of prediction. *J. clin. Psychol.*, 1950, 6, 387–392.
10. Terman, L. M., & Merrill, M. A. *Measuring intelligence.* Boston: Houghton Mifflin, 1937.
11. Wechsler, D. *The measurement of adult intelligence.* (3rd Ed.) Baltimore: Williams and Wilkins, 1944.

Received February 9, 1951.

THE FORMATION OF SOCIAL GROUPS UNDER CONDITIONS OF SUCCESS AND FAILURE[1]

J. C. GILCHRIST
University of Wisconsin

Various sources indicate that a success-failure continuum is an effective variable in the determination of group formation (1, 2, 10, 11, 16, 17, 18, 19, 20, 21, 22). The present study is an attempt to investigate one aspect of the problem in a laboratory situation. Precisely, the problem is as follows:

Given:

1. a success-failure continuum the definition of which is held in common by each of the participant individuals,
2. that these participant individuals are motivated to perform in terms of the tasks involved in that continuum,
3. that these individuals are distributed along the success-failure continuum so that, working as individuals, some of them experience success and others failure,
4. that objective success and failure occur in direct relationship with subjective success and failure,
5. that these participant individuals exist cognitively for each other, and
6. that every individual's "objective" (commonly defined) success or failure be communicated to every other participant individual,

what will be the grouping choices of a random collection of individuals when action not of individuals but of groups is required for further possible goal attainment?

And, if the above grouping choices are made operative in various ways and the groupings so formed subjected to the further operation of the success-failure continuum,[2] what will occur in terms of further grouping choices?

The problem, however, is more extensive than an attempt to answer the above questions on the basis of experimentally gathered data. It is also an attempt to derive the consequences of the application of success and failure upon grouping choices through the use of two hypothetical constructs of psychology. These constructs are (a) level of aspiration and (b) expectation. The limitations and definitions given respectively for these variables by Lewin, *et al.* (6) and Tolman (14) are accepted herein.

The Predictions

Prediction 1. S's[3] will choose other S's for group activity along the same or similar lines of endeavor as that on which they were labeled.

Prediction 2. F's will choose S's for group activity along the same or similar lines of endeavor as that on which they were labeled.

Under the conditions given in the formulation of the problem and where the task involved is (a) familiar and/or of a class subsumed under a frame of reference previously established by the individual and (b) perceived as neither too difficult nor too easy by the individual, he will establish a level of aspiration with respect to that task. If, then, some persons working as individuals on the same task as the other participating individuals exceed or equal their level of aspiration, they will experience success thereby establishing an expectation of future personal success when working on the same or similar tasks. And, if their success is communicated to other participants working with the task —at least in the sense that they have "objectively" attained the goal—the other participants will label them successful, i.e., establish an expectation that these persons will succeed

[1] This study is part of a thesis submitted in 1950 in partial fulfillment of the requirements for the Ph.D. degree at the University of California at Los Angeles. The writer wishes to thank Drs. Franklin Fearing, J. A. Gengerelli, Walter Goldschmidt, and Harry Hoijer for their valuable criticisms and assistance.

[2] In the remainder of the discussion, the success-failure continuum will be treated in its simplest fashion; that is, as a dichotomous variable.

[3] Hereinafter, the letter "S" will stand for those labeled successful and the letter "F" for failure. Sequences of these letters will indicate sequences of success and failure on tasks. Thus, an SF will be an individual who succeeded and then failed while an FF is one who has twice failed.

in the future when working on the same or similar tasks. On the other hand, if some participants have performances below their level of aspiration they will experience failure thereby establishing an expectation of future personal failure when working on the same or similar tasks. Here, too, if their failure to attain the goal is communicated to the other participants in the task, these other participants will label them as failing, i.e., establish expectations that these failing persons will fail in the future when working on the same or similar tasks.

The above considerations lead to the empirically substantiated statement that the subjectively successful persons will tend to increase their previous level of aspiration while the failure persons will lower theirs for continued activity along the same dimensions as the previous task (see [6]).

If at this point further behavior along the same or a similar line of endeavor is necessary except that the concerted action of two or more individuals is required, there will exist a pool of individuals, some successful and some failures, from among whom choices of work partners may be made. When an S is choosing he will be operating in terms of (a) a need for a successful choice and (b) a need to maintain and attain his now raised level of aspiration. No expectation as to the success of any given choice exists and, therefore, the choice situation is itself ambiguous. But working with an individual labeled a "success" would support the expectation of attaining the S's raised level of aspiration whereas working with a "failure" would bring a reduction in the aspiration level and/or reduce the expectation of attaining it. Hence, in the absence of structured choice expectations, S's will choose each other.[4]

For the failures, the conditions which exist are (a) the unsatisfied need for commerce with the goal as a result of the uncompleted first task, (b) the need for a successful choice, and (c) the need of at least maintaining and attaining their now lowered levels of aspiration. As in the case of the S's, the possibility of a successful choice is an unstructured situation, but the direction of choice is as clearly indicated. Expectation of future success with the other participants is delineated by their success or failure labels, and, as before, a successful choice of an S would lead to an increased expectation of future task success. Hence, F's will choose S's.

Prediction 3. If S's are grouped (i.e., permitted the operation of their choices as predicted in 1 supra) and are successful when working together (SS), they will, in terms of a second grouping choice, choose one another.

Prediction 4. If F's are forced to group with F's (i.e., not permitted the operation of their choices as predicted in 2 supra) and fail when working together (FF), they will, in terms of a second grouping choice, distribute their choices between the SS's and their fellow FF's.

When the operation of the reciprocal choices of the S's is allowed in a fashion which denies the F's choices of the S's, then the S's have again achieved a goal through a successful choice. If the S's then succeed when working together, they will choose one another for the same reasons as advanced in the support of Prediction 1. Now, however, two additional factors support the prediction that SS's will choose each other: (a) The successful first choice will have established expectations of a second successful choice if they again choose each other. (b) The experience of goal achievement while working together will have resulted in increased levels of aspiration and heightened expectation of success for working together.

At the same time, if the F's fail working together, they have been subjected to three nonachievement experiences: failure as individuals, failure in their choices, and failure working together. These facts lead to conflicting vectors within the FF's: for future task success they should choose from among the SS's but for choice success they should choose from among their fellow FF's. This latter statement follows from the fact that the S's had worked as a unit to become SS's and the F's as a unit to become FF's. Thus, their expectation will be that these groupings

[4] The operation of frames of reference peripheral to the attainment of the task goal (e.g., physical appearance standards, clothing standards, etc.) may enter into the choice in such a fashion as to negate this and the following predictions to some extent in a completely free situation. In the event of a relatively homogeneous population and/or where the task is sufficiently important to the participants, however, the prediction should stand.

will recur. Therefore, FF's to obtain a successful choice should choose other FF's. This ambivalent condition will be reflected by the choices of the FF's: taken as a whole, the pool of FF's will choose neither the SS's nor the FF's significantly one over the other.[5]

Prediction 5. If S's are grouped and fail when acting together (SF), they will, on a second grouping choice, choose their fellow SF's.

Prediction 6. If F's are forced to group with F's and are successful when working together (FS), they will, in terms of a second grouping choice, choose their fellow FS's.

The SF's will have undergone a success experience as an individual, a successful choice, and a failure experience when working together. Under these conditions, the SF's will have not attained the goal and thus will lower their level of aspiration for future performance with their fellow SF's. This factor alone would indicate that the now SF should choose an FS, but the expectation with respect to the FS's is in itself ambiguous; they have failed as individuals and then succeeded when working with other initial F's. Besides, other considerations enter here. Firstly, the expectation that a second successful grouping choice will occur if the same grouping choice is repeated. Secondly, a single failure following two successes should be less effective in bringing about a change in behavior than if it had not been so preceded. The latter statement follows from the repeated finding (4, ch. 8) that reinforcement (reward, success) of a response increases the resistance of that response to extinction. For the SF's, task success expectations are ambiguous, but choice success expectations center upon the other SF's. Hence the prediction that, although the SF's have failed working together, their choices should be of their fellow SF's.

On the other hand, the FS's have experienced failure on the first task, a failing first choice, and a success experience when working together. Thus, together, for the first time in the series they have achieved a goal. This experience should be more intense than one not preceded by two failures since this goal attainment will also serve as substitute satisfaction for the two previous unsatisfied needs.[6] This effect should produce emphasized expectation of success in future activities when working together. Additionally, the original S's are now SF's and, thus, expectations of task success with them is ambiguous. Further, expectations of choice success have been established upon their fellow FS's. These considerations lead to the prediction that the FS's will choose each other.

Prediction 7. If S's are forced to group with F's, then, on the basis of there being two such mixed groups and that one of these mixed groups succeeds while the other fails,

(a) the SS's of the SS and FS group will distribute their choices between the SF's and FS's while

(b) the SF's of the SF and FF group will distribute their choices between the SS's and the FF's.

Prediction 8. If F's are grouped with S's, then, on the basis of there being two such mixed groups and that one of these mixed groups succeeds while the other fails,

(a) the FS's of the SS and FS group will choose the SS's with whom they have worked while

(b) the FF's of the SF and FF group will distribute their choices between the SS's and SF's.

At this point, the situation becomes increasingly complex, for, when the reciprocal choices of the S's for each other are prohibited and they are forced to group with F's thus making the F's choices of S's successful,[7] the result is at least two groups composed of at least one S and one F each. If, in the mini-

[5] It might be assumed that the intensity of a success or failure experience following a grouping choice is initially less than that of the task since the goal of the behavior sequence is task success. The whole question of intensity relations between task and grouping choice success and failure as well as its relation to repeated experiences of success and failure is one which needs investigation. In this investigation, a constant intensity sufficient only to make the variables operative is assumed.

[6] Lissner (8) found that the need residual from uncompleted tasks was "discharged" by the subsequent completion of another task as measured by frequency of resumption of the uncompleted tasks.

[7] It should be pointed out that for this experiment this is a somewhat inaccurate statement. In terms of the experimental design it would sometimes occur that on the first choice both F's would choose the same S. Thus, at least some of the FS's and FF's did not get to work with the particular S whom they had chosen. Hence, in this section a successful first choice is considered as being paired with either S.

mum instance of two work units, one of them is made to succeed and the other to fail, the further result is one unit composed of an SS and an FS and another unit composed of an SF and an FF. These persons will have the characteristics of their counterparts discussed in Predictions 3 to 6 except that those initially labeled F's have a successful first choice while those labeled S have a failing first choice. In this situation, the SS would have only one residual tension system in the series of events, the failing first choice, which would have been subjected to some substitute satisfaction following success on the second task. Task success expectations indicate a choice of either the SF or the FS although precedence must be given to the FS since it was with him that the SS has just succeeded. Choice success expectations indicate a choice of the FS, for it was with him that the SS was grouped. From the above considerations would come the prediction that choices of the FF by the SS will not occur and that the SS will choose the FS. Another factor, though, makes this prediction ambiguous. The SS, in view of his two successful performances, the last of which resulted in at least a partial substitute satisfaction for the failing choice, may not have sufficient motivation at the requirement of a second choice to be affected by the expectation of its failing, and hence he may act as if there was no need to shift his choice from the now SF to the FS. This last fact and the task success ambiguity of the FS and SF require that the prediction be made only in terms of the SS's distributing their choices between the SF's and FS's.

The SF, on the other hand, when presented with the requirement of a second choice, will have the need thus established increased by the unsatisfied needs residual after his failing first choice and second task and his motivation, therefore, should be relatively high. The expectation for a successful future performance will be in terms of either the SS or the FS. However, the expectation of task success with the SS will be higher since he has twice succeeded whereas the FS has been successful only once and that while working with the SS. This factor indicates that the SF will continue in his choice of the now SS's for task success but, in terms of choice success, the SF's expectation is that, if he continues in his choice of the SS, his choice will again fail. Hence, for a successful choice, the SF has the possibilities of either the FS or the FF. It was with the FF, however, that the SF was grouped and that will have established an expectation that such will again take place. These two competing vectors lead to the prediction of the distribution of choices by the SF's between the SS's and FF's.

As to the FS at the requirement of a second choice, his course is clear. His expectations of future task success and choice success are both anchored to the now SS's and this dictates the prediction that the FS's will choose the SS's.

The FF, in this case, is the only individual who has not had a successful task experience although he has made a successful choice. With respect to task aspiration, they can best be maintained or increased by choice of the SS, SF, or FS with emphasis upon the SS since he is the only person doubly labeled a success. It is with a choice of the SF, however, that the FF will have greater expectations of his choice being successful for it was previously successful. Hence, the prediction that the FF's will distribute their choices between the SS's and the SF's.

Prediction 9. The same tendency of grouping choices as predicted in 1 through 8 supra will spread to other activities. This spread will be effective in inverse proportion to the degree of importance which these other activities have for the individuals' other frames of reference.

Every phenotypical definition of the success-failure continuum will have limiting cases. These limitations will be a function of the other frames of reference of the individual. Similarity of the frames of reference and the value placed upon the contemplated action by the other frames will be the major determiners of these limitations. Where the other contemplated activity is similar to that along which a specific success-failure definition was made and provided the individual has no other previously established reference scale for the specific persons involved, the labeling of success or failure will generalize so as to affect the choices for group action in the other activities in the same direction as that postulated in Predictions 1 to 8 above.

As the degree of similarity decreases, this generalization effect will lessen and in proportion to the degree of importance the future action has for these dissimilar frames. The first of these statements follows from the principle of stimulus generalization in the learning process. The second follows from the finding (5) that unless the activity is valued (taps an ego-involved frame of reference) little or no motivation may be inferred. In the absence of motivation, there will be little shift in the perception of the situation, and, consequently, little change in the previous choices.

IMPLEMENTATION OF THE PREDICTIONS

In order to test the predictions, four experimental groups of subjects were required. Since Predictions 1 and 2 bear upon the grouping choices after the first task performance, the data from all subjects could be used to test these predictions providing the four experimental groups were treated exactly alike on the first task and choice. This was done. Four subjects were brought into the laboratory at a time, and two of these were made to succeed in the performance of a task as individuals while the other two were made to fail. This individual success or failure was communicated to the participants, and, after the time limit for the task had elapsed, each subject made a grouping choice of one of the other participants for continued activity on a similar task to be performed in pairs. Beyond this point, however, the treatment of the subjects varied. Some collections of four subjects had their sessions completed by answering a set of sociometric questions. This set of subjects was Group I and was obtained in order to test generalization of grouping choices (Prediction 9) following the single success-failure experience. In Group II, the two S's were paired and made to succeed while working together on a second and similar task, while the two F's were paired and made to fail. Following this, another grouping choice was made, and then the subjects made the other sociometric choices (Predictions 3, 4, and 9). Group III was treated in the same fashion as Group II except that the S's failed when working together while the two F's succeeded (Predictions 5, 6, and 9). In Group IV, each S was paired with one of the F's and one pair made to succeed while the other failed. As in Groups II and III, the success and failure as pairs was communicated to the participants and, after the second grouping choice, the subjects made the other sociometric choices (Predictions 7, 8, and 9).

Another group, Group Ia, was treated in the same fashion as Group I, but only three subjects participated in a given experimental session and only one of the three was allowed to succeed.

Since the dependent variable, grouping choices, is enumeration data, it permits the use of the small sample statistic, chi-square, with the expected frequencies (null hypothesis) of choice of either an S or an F determined by the proportions of S's and F's in the pool from which a choice is to be made. Hence, in Prediction 1 when an S was choosing, he had a chance value of one-third for choosing an S and two-thirds for choosing an F. The reverse was the case when an F was choosing (Prediction 2). Similar values prevail for all other predictions except 7, 8, and 9. In Predictions 7 and 8, the values would be one-third for any choice since all four of the participants were differently labeled. In Prediction 9, the values were the same as those of the particular prediction under whose conditions it was tested.

The size of the N required dictated that college students had to be used as subjects because they were the only relatively homogeneous and accessible pool of adult subjects of that size available.

The source of subjects having been determined, the problem became one of making the experimental model as much like a "real" situation as possible for the subjects. With college students it was decided that a test development program would be a meaningful guise, and that "reasoning" tests would furnish a frame of reference in terms of which the initial level of aspiration would be established. Further, besides the motivation provided by the establishment of a level of aspiration on the "reasoning" test, it was postulated that if they expected that their success or failure were made a matter of open knowledge to the other participants an ego-threat would be presented and furnish heightened need for success as well as tapping the competitive frames of reference which both the general American cultural setting and the university situation inculcates. Under this motivation-arousing setup, zero levels of aspiration should be rare among the members of a fairly select student body. Therefore, if interesting test materials were selected which were capable of being manipulated by the experimenter in such a manner as to produce either complete (or almost complete) objective success or complete (or almost complete) objective failure, S's should have positive attainment discrepancies and F's should have negative attainment discrepancies.

The first of these tests was that originated by Rashkis, Cushman, and Landis (12). This test calls for the double classification on a board arranged as a 4×4 table of the following words or phrases: THIRST, DRINKING, LIQUID, FOUNTAIN, HUNGER, EATING, FOOD, RESTAURANT, NEED MONEY, WORKING, JOB, INDUSTRY, SEX DESIRE, SEX ACT, WOMAN, MARRIAGE.

In the modified form used in this experiment each word or phrase was typewritten on one of 16 1-inch cardboard squares and placed in a small plain blue envelope. Four such sets were provided; one for each of the participants at a single experimental session. This material combined with the ambiguous instructions given the subjects ("Do with the cards whatever you think you should do.") provided a situation where any of a number of solutions could be called "correct" or, inasmuch as there were so many possibilities for rearrangement, any solution could be called "incorrect."

The second test to be modified was the Vigotsky concept formation test as described by Hanfmann and Kasanin (3). The original test consisted of 22 wooden blocks in 5 different colors, 6 different shapes, 2 heights, and 2 sizes of the upper surface. On the under side of each block was written one of four syllables: *lag, bik, cev,* and *mur. Lag* was on the tall-large blocks, *bik* on the short-large blocks, *mur* on the tall-small blocks, and *cev* on the short-small blocks. The purpose of the test was that the testee segregate the blocks bearing each syllable into a separate group. As modified here one set of 15 wooden blocks was constructed classifiable either on the basis of short curvilinear base blocks, tall curvilinear base blocks, short rectilinear base blocks, and tall rectilinear base blocks or on the basis of finding slight but discriminable height differences between the four kinds of blocks. A second set of blocks was constructed in similar fashion except that the syllables were assigned to these blocks at random and three additional blocks destroying the height relationship and having both curvilinear and rectilinear edges on their bases were placed in the set. The first of these sets was constructed to make the induction of success relatively simple; the last was, of course, the failure set. Each set was placed upon a covered flat wooden tray and administered under the Vigotsky instructions.

Other requirements of the test setup were (a) the prevention of subjects from finding out that they were being differentially treated by the experimenter and (b) the communication of success and failure of the subjects. To accomplish the first of these, in the middle of the otherwise bare experimental room, a large table and four chairs were placed. The top of this table was divided into equal areas by an arrangement of partitions and blackout cloth. In this manner, a private working space was provided for each subject in such a way that he could easily see the wall to his front and the experimenter standing in the area on the other side of the table.

The second factor, that of communication of success and failure of the subjects, was accomplished by several means. As each subject entered the room he was given a neckcloth which had a panel bearing a number from one to four on both the front and the back, and the working spaces at the table were also numbered in order from one to four. The subject was to be seated at the work space bearing the same number as that of his neckcloth. On a shelf 7 feet above the floor and against the wall which the subjects faced were an easily seen electric timer and a signal panel having individually illuminated numerals from 1 to 4. To justify announcement of a success, a time limit was imposed on the "test" situation. "You will have as much time as it takes the first two persons (On the second task, 'pair' was substituted for 'two persons') to complete the test plus three minutes. . . .

"When someone completes ('when the first pair') the test, I will announce it and turn on the numbered signal light which corresponds to the number he wears ('numbers they wear')." (It was thought that perhaps this unusual time limit might arouse suspicion on the part of the subjects that all was not as it appeared. Only a few subjects gave any indication that it did, and the data of these subjects were discarded.)[8]

Beyond these conditions, the operation of peripheral frames of reference had to be blocked so that only the success-failure continuum in the task dimension was operative. To this end, the four subjects were, at any one of the experimental sessions, either all males or females. Additionally, the factor of acquaintanceship of any of the subjects was controlled by asking at the end of the session if any of the participants had been previously acquainted, and, if any answers were "yes" their data were discarded. In order to prevent acquaintanceship arising between the subjects while they were awaiting the start of the experiment, a long biographical questionnaire was constructed for the purpose of preventing any conversation between the subjects prior to their working in pairs.

The possibility of there being position preferences in choice was controlled by randomly selecting prior to the experiment the numbers of the pair to succeed at every session.

Grouping choices were recorded by the subjects upon sheets of yellow copy paper in response to the statement "You are going to be retested in just a moment on a similar test working together in pairs. Please take the yellow sheet of paper in front of you and write down the number of the person with whom you wish to be paired." Each subject was provided with a separate and new sheet for each choice.

To test Prediction 9 a social questionnaire was prepared calling for grouping choices in two different hypothetical situations; one similar to that of the test and one dissimilar. The questions were:

1. To which one of the persons taking this test will you be most likely to speak when next you see him (her)?

[8] A control experiment was run to determine what effect, if any, the announcing and signalling of the numbers of two of the four subjects had upon grouping choices. Twelve sets of four subjects each heard announced and saw signalled the numbers of two subjects with no explanation given as to the reason for so doing. Six sets were all males and six sets were all females. A different one of the six combinations of the numbers 1, 2, 3 and 4 taken two at a time was announced in each of the six male sets and six female sets. The subjects then made a grouping choice under the same requirements as in the experimental sessions. The results were that the persons whose numbers were announced chose unannounced persons in 20 instances and announced persons in 4. Chi-square equals 2.30 which has a p of between .20 and .10. The persons whose numbers were not announced chose unannounced persons in 14 instances and announced persons in 10. Chi-square equals 5.67 which has a p of less than .02 with one degree of freedom. Because the direction of choice in both cases is opposite to that predicted under success and failure conditions, it is held that if the predictions hold, the result cannot be explained upon the basis of announcement and signalling of numbers per se. The significant value of chi-square obtained from the choices made by unannounced persons cannot at present be satisfactorily explained.

2. Which one would you most prefer asking to the Co-op [9] for a cup of coffee or a coke?
3. With which one would you most prefer to go on a "double date"?
4. With which one would you most prefer to work on "math" problems?
5. With which one would you most prefer to study for examinations?

primarily from psychology classes at the University of California, Los Angeles, during the summer sessions of 1948. One hundred and twenty-six of the subjects were males and 108 were females. They ranged in age from 15 to over 50 years of age with the median occurring within the 20–24 years interval. They reported that their years of college ranged from zero to over five with the median in the vicinity of three years and that their college grades averaged from "D" to "A" with the median at "B."

TABLE 1

THE FREQUENCIES OF FIRST AND SECOND TASK GROUPING CHOICES *

GROUP	CHOOSER	FIRST CHOICES OF		SECOND CHOICES **			CHI-SQUARE OF SECOND CHOICES	*df*	*p*<
		S's	F's						
I	S	15	5						
	F	15	5						
II	SS	14	6	SS 18		FF 2	26.18	1	.001
	FF	16	4	SS 11		FF 9	0.72	1	.50
III	SF	16	4	SF 16		FS 4	17.38	1	.001
	FS	15	5	SF 6		FS 14	10.38	1	.01
IV	SS	16	2	SF 11	FS 6	FF 1	8.34	2	.02
	SF	15	3	SS 10	FS 2	FF 6	5.34	2	.10
	FS	14	4	SS 12	SF 4	FF 2	9.34	2	.01
	FF	11	7	SS 7	SF 8	FS 3	1.56	2	.30
Sum of initial S's choices		76	20	Chi-square=90.8; $p<.001$					
Sum of initial F's choices		71	25	Chi-square=2.30; $p<.20$					

* Yates' correction has been applied in every case where an expected frequency was less than ten and $df=1$.
** The task labels of the chosen individuals are shown in the separate cells.

The first three questions were, of course, non-academic questions and were selected so as to be in a rough three-step scale of importance to the rating-dating complex of college students' frames of reference as shown by Waller (15) and Loomis and Green (9). These questions were asked orally and the answers were recorded by the subjects. An answer to the questions was not forced by explicit directions on the first reading but was upon a second.

Subjects

The volunteer subjects whose data were used in this experiment numbered 234. They were drawn

[9] "Co-op" is the local term for the campus cafeteria and coffee shop.

RESULTS AND DISCUSSION

Predictions 1 and 2 bear solely upon the first choices of the subjects, and, inasmuch as all groups except Ia had the same number of subjects and were treated in exactly the same fashion until after the first choices were made, their data were combined. The data are in Table 1. It will be noted that Prediction 1 is supported.

Prediction 2 does not fare so well under

the same analysis. F's choose S's primarily as predicted (see Table 1) but chi-square falls between the 20 and 10% levels of confidence. It has been thought that this situation might occur since it seemed likely that the parameter might, in this experimental condition, fall close to the chance value. Therefore, Group Ia was obtained. As might be expected, the requirement that four persons keep appointments at the same time and place resulted in a number of instances when only three persons appeared. It was decided to use the same procedures when three rather than four persons appeared except that only one subject was made successful. In this case, an F making a grouping choice would have one S and one F between whom to select. Thus, the null hypothesis expectation would be that one-half of the choices be of F's and one-half of S's. The question asked was "If you were to work on a similar test in pairs, with which one of the other persons taking this test here today would you rather be paired?" This less real question casts doubt upon the result since it is somewhat in line with the complaint that laboratory investigations in the social sciences are artificial. The fault is admitted to some extent, but the result is held to indicate, at least, the validity of Prediction 2. The result was that 20 of the F subjects chose S's and 7 chose other F's.[10] This result yields a chi-square of 6.26 which with one degree of freedom is significant beyond the 2% level of confidence. Although the confidence which can be placed in Prediction 2 is not as great as that of Prediction 1, it would seem that both are supported by the data. Evidently when a random collection of individuals is required to form groups, the individuals tend to form them around the successful individuals. Failures do not form a differential group (i.e., reciprocate one another's choices), then, on the basis of their mutual failure. The emergence of such groups must be determined by events occurring after the initial grouping choice has been made.

These other events are obviously the erection or occurrence of some kinds of barriers to the operation of the grouping choices. Such barriers may exist for a number of reasons. Since the S's typically tend to choose each other reciprocally, if the external situation requires that two or more persons act together and bars more than just the number of S's from operating together, the F's are necessarily thrown together. Or, if the S's actively reject the participation of the F's, the same condition results. These and related problems remain for future investigations, but Predictions 3–6 bear upon the general question of any barrier at all existing to the operation of the typical choices of the F's.

Predictions 3 and 4 deal with the case of continued success of the S's and continued failure of the F's. Prediction 3 was that the SS's would choose each other reciprocally and Prediction 4 was that the FF's would distribute their choices between the SS's and their fellow FF's. Table 1 shows that this is the case. There is one set of reciprocated choices: the SS's are choosing each other. The FF's, on the other hand, demonstrate no clear-cut set of choices either reciprocal or of the SS's. This state of affairs leads to the suspicion that the failures will not form a group until all possible expectations for the occurrence of a successful choice of a successful individual have been exhausted. Even then, it is suspected that the group so formed would not be stable; that any indication that there could be a possible successful choice of a successful individual would result in the dissolution of such a group. These latter statements also remain for future study, but there is some evidence hinting at their correctness.

The questionnaire dealing with sociometric choices in other situations had two questions set up as tapping a frame of reference similar to that of the tasks: "With which one (of the persons taking this test) would you most prefer to work on 'math' problems?" and "With which one would you most prefer to study for examinations?" These questions permit the respondent to behave autistically; they permit the choice to be made in terms of preferences without the necessity of these preferences being tested. Thus, by comparing the direction of response to these questions to the responses made in the more real situation (i.e., the second grouping choice condition), an estimate as to what would occur if the FF's perception of the situation was changed to provide any expectation that a

[10] The choice made by one subject was uninterpretable and had to be discarded.

successful choice of a successful individual could—or was to—occur.

It will be remembered that, on the second grouping choice, FF's chose SS's 11 to 9 over their fellow FF's. On Question 4 of the Social Questionnaire, however, the FF's chose the SS's 19 to 1 over the other FF's while on Question 5 the choices favored the SS's 16 to 4. Although the experiment was not designed to test the significance of changes between a first and second set of responses and therefore the N is too small for such a chi-square check to be conclusive,[11] it can at least furnish an indication as to the possible correctness of the hypothesis. The chi-square for the changes between the second grouping choice and the response to Social Question 4 is 6.13 ($p<.02$) while that of the second grouping choice on Social Question 5 is 1.78 ($p<.20$). Both the changes are in the expected direction: the social questions produced more choices of the SS's than did the second grouping choice. In view of the small N and the last value of p, it is only tentatively suggested that groups which might form under continuing task and choice failure are unstable.

Returning to the narrower fields which the experiment was designed to test, Prediction 5 was that SF's would choose SF's, and Prediction 6 stated that FS's would choose each other. The predictions are substantiated by the data as shown in Table 1.

The results of Group III with which Predictions 5 and 6 were tested lead to the conclusion that differential group formation is here emerging. The SF's are choosing each other and the FS's are also choosing each other. Thus, in the face of one failure experience when working together the initially successful persons still tend to form a grouping, but after the occurrence of a success experience when working together the initially failing persons tend to form a distinct positive bias towards one another. It would appear, then, that definite establishment of groupings by initially labeled failure individuals is dependent upon the attainment of some measure of task success by those persons unless, as was pointed out in discussing the results under Prediction 4, any possible expectations of a successful choice of a successful person have been negated either through inference or direct experience.

Group IV was the group designed to test Predictions 7a, b and 8a, b. It will be recalled that the initially labeled S's were not permitted the operation of their reciprocal choices thereby making the F's choices of S's successful. Prediction 7a stated that under these conditions the SS's would choose in terms of the SF's and FS's and ignore the FF's. The data are presented in Table 1. When the choices of the SF's and FS's by the SS's are combined and a chi-square computed with the other cell composed of the SS's choices of the FF's, one degree of freedom results which permits the use of Yates' correction. The result is a chi-square of 5.07 ($p<.05$) which indicates that the FF's are ignored by the SS's, but the main contribution to the size of chi-square arises from the piling up of choices upon the SF's. Thus, that portion of the prediction which stated that the SS's choices should be distributed between the SF's and FS's is not supported. Rather, it would appear that the SS's were primarily choosing SF's and ignoring the FF's.[12] Several explanations for this predictive failure exist: the SS's after their two task successes may not have sufficient motivation resulting from the requirement of a second choice, for their success on the second task should result in a substitute release of tension for the failing first choice. It may also indicate that the initial S and F labeling of the subjects furnishes stronger expectations of success and failure with the persons so labeled than does the subsequent task labeling.[13] This condition may arise from the fact that the subsequent labeling is a function of group action and not solely a function of the individual.[14] If either or both of these

[11] Lewis and Burke (7) in a recent review of the misuses of chi-square have stated that an expected frequency "of 5 is believed too low." There are, of course, differences of opinion upon the matter. Snedecor (13, p. 169) reports an unpublished study by Fryer in which "the small expected frequencies of mutations in radiation experiments, as low as 0.5" were examined. "In a series of 47 tests . . . no conclusions about significance . . . had to be changed when the more difficult exact tests were made." The lowest expected frequency here is 2.8.

[12] It should be noted, however, that 5 of the 6 cases choosing FS's had originally chosen the now SF's.

[13] If this is the case, it may account in part for the continued choices of each other by the SF's of Group III. It would also account for the lack of a higher number of reciprocated choices of the FS's of the same group.

[14] The formulation that the initial reinforcement, according to Hull (4, Ch. 8) furnishes the greatest

situations were operative, it would account for the lack of change in choices by the SS's and thus having the frequency of FS choices at that which might be expected by chance alone.

Prediction 7b stated that the SF's would choose SS's and FF's to the exclusion of the FS's. Table 1 shows a chi-square with a $p<.10$. When the choices of SS's and FF's by the SF's are combined with the choices of FS's completing the other cell, chi-square equals 3.06 which also exceeds the 10% level of confidence. The direction of events is as predicted except that the major contributions to chi-square arise from the SF's choices of the SS's and ignoring the FS's, but the confidence to be placed in the prediction is not high. The occurrence of a high number of choices of SS's here may again indicate that, inasmuch as the initial goal of the behavior sequence is apparently task success, it is in the early stages of failure more important than choice success.

Prediction 8a stated that the FS's of Group IV would choose SS's to the exclusion of the SF's and FF's. The data sustain the prediction as Table 1 shows. When the SF and FF choices are combined to form one cell and a chi-square computed with SS choices forming the other cell, the result is a value of 7.56 ($p<.01$).

Prediction 8b stated that the FF's of Group IV would distribute their choices between the SS's and SF's and ignore the FS's. The data are in Table 1. When the SS and SF choices are combined into a single cell, chi-square is 1.56 which has a $p<.30$. The direction is as predicted although statistically insignificant.

As far as Group IV is concerned, one thing is apparent from the data: differential group formation did not clearly occur. There were no notable reciprocal choice tendencies between members of the same work unit nor between initial S's nor between initial F's. The SS's were primarily choosing SF's and ignoring FF's, the SF's were primarily but insignificantly choosing SS's and ignoring FS's, the FS's were primarily choosing SS's and ignoring FF's, and the FF's were choosing no one primarily. The predictions, except for that regarding the FS's, do not stand up in detail although generally they are in the right direction. It will be noted, however, that the detailed failure of the predictions occurred where a distribution of choices had been expected. Beyond the explanations already advanced, two explanatory possibilities intrinsic in the experimental situation appear plausible. (a) The efficiency of the specific setup in producing motivation, subjective success and/or failure, and communication may have been insufficiently high so that when applied to this last and more demanding predictive situation the design failed, and (b) the parameters of choice under conflicting task and choice success expectation may lie at or near the chance values chosen for this experiment. These factors, in addition to the added postulations which seem to be dictated by the earlier findings, make it seem uneconomical to continue to test the details of Predictions 7a, b and 8b by the current design.[15]

Prediction 9 stated that the grouping choices made on the tasks would generalize to other activities in direct proportion to the similarity between the other activities and the activity along which the success and failure labeling occurred. It further stated that such generalization would take place in inverse proportion to the importance of these other activities to other frames of reference. As initially made, the prediction was, of course, contingent upon Predictions 1 through 8 being upheld. It has been seen that Predictions 1 through 6 and 8a were sustained but that 7a and 7b and 8b required the introduction of new variables to explain the results. Prediction 9, therefore, must be changed in that part pertaining to Predictions 7a and 7b to conform with the results obtained while that part pertaining to 8b remains equivocal. But other factors enter before the data can be considered.

It has already been pointed out that the social questions permitted the respondent to behave autistically. From this fact, it is argued that, wherever task and choice success expectations conflict regarding the predicted direction of choice, such questions

increment of habit strength also may account for this latter point if expectation of task success established via communication and direct experience be considered as following the same growth function as habit strength in the presence and absence of reinforcement.

[15] Another test is currently being made.

when asked along dimensions similar to the task will result in the autistic operation of task success expectations. Further, questions about choices for dating activities will result in the autistic operation of choice success expectations. Hence, for that reason, data obtained under conditions of conflicting choice and task success expectations should not be considered a test of the prediction when the prediction rests upon the conflict of task and choice expectations of success. This condition would exclude from consideration the FF's of Group II, the FS's of Group III [16] and the SS's, SF's, and FF's of Group IV. Table 2 presents the data for all the labeled individuals in all the groups, but the following discussion excludes those listed above and the F's of Group I for the same reasons as stated earlier.

One further point, however, has to be settled. Are Questions 1, 2, and 3 tapping a different frame of reference from Questions 4 and 5? If a greater number of refusals to answer Questions 1, 2, and 3 than Questions 4 and 5 on the first reading be accepted as proof, then the answer is "yes," for there were 56 refusals on Question 1, 65 on Question 2, 77 on Question 3, 44 on Question 4, and 50 on Question 5. The chi-square of refusals on 2 and 3 versus those on 4 and 5 is significant at 1% level of confidence.[17]

Turning to the data in Table 2, it is found that, of the 5 sets of labeled individuals under consideration, there are 25 possible points at which the directions of the predicted generalization of the success-failure labeling can be tested: once for each of the five sets on each of the five questions. Of these 25 check points, 24 follow the predicted direction and one is questionable. That one is the result of the FS's of Group IV on Question 3, the most important one of those relating to the rating-dating frame of reference. It would appear that generalization does occur. And the spread to other activities is above the 5% level of confidence in 21 of the 25 cases. But is the generalization effective in inverse proportion to the importance these other activities have for other frames of reference? To answer this question, the distribution of the significant spreads among the various social questions must be examined.

For the five sets of data under examination, there are 4 insignificant spreads out of 25. Of these 4 insignificant spreads, 3 of them occur on the question relating to "double dating" (Question 3), the most valued of the three activities touching the rating-dating frame of reference, and the remaining insignificant spread occurs on the question relating to studying for examinations, the most important of the two activities in the academic frame of reference. The evidence cannot be considered conclusive but is indicative that generalization does occur until it taps an important aspect of another frame of reference.

The results—as far as the experiment per se is concerned—are generally clear. The interpretation of these results for understanding and predicting everyday social behavior, however, requires consideration.

In everyday social behavior, the results of this experiment should be most applicable in situations involving small, face-to-face collections of individuals where the individuals concerned provide new and ambiguous or unstructured means and/or goal objects for each other. Such a situation would occur (a) when the individuals had never previously interacted together or had never previously observed each other in an interaction situation or (b) when previously interacted-with individuals are to-be-operated-with in drastically different means-end fields.[18] These conditions are limited by the total societal setting. One possible way of conceptualizing this

[16] Under the new assumption of greater strength of expectation being established by the initial labeling of S or F, these FS's would be expected to express that greater strength of expectation autistically on the task question.

[17] Question 1 was excluded from the computation since it was believed it constituted a minor factor in the rating-dating complex. A questionable assumption of independence of answers also enters in this chi-square.

[18] Human beings, at least, never will completely fulfill these conditions. Under condition (a), the individual would bring and generalize to the new individuals various frames of reference (e.g., physiognomy, skin color, dress, speech, etc.) which would in part determine his expectations with respect to those individuals. Under condition (b), it would be expected that somewhat the same qualifications would apply. There are hardly any conceivable new and drastically changed means-end situations which would not have some backward time reference through other and/or similar frames of reference with respect to the other participating individuals. Such conditions as (a) and (b) above may be stated only relatively.

TABLE 2

THE FREQUENCIES OF FORCED GROUPING CHOICES ON THE ITEMS OF THE SOCIAL QUESTIONNAIRE *
(The task labels of the chosen individuals are shown in the separate cells.)

Group	Chooser	Question Number																			
		1				2				3				4				5			
				$p<$	**			$p<$	**			$p<$	**			$p<$	**			$p<$	**
I	S	S 12	F 8	.05	+	S 16	F 4	.001	+	S 10	F 10	.20	+	S 16	F 4	.001	+	S 19	F 1	.001	+
	F	S 14	F 6	.95	+	S 13	F 6	.95	−	S 14	F 5	.95	+	S 12	F 8	.80	−	S 12	F 8	.80	−
I_a	F	S 20	F 7	.02	+	S 20	F 6	.01	+	S 15	F 11	.50	+	S 20	F 7	.02	+	S 16	F 10	.50	+
II	SS	SS 17	FF 3	.001	+	SS 17	FF 3	.001	+	SS 16	FF 4	.001	+	SS 19	FF 1	.001	+	SS 19	FF 1	.001	+
	FF	SS 10	FF 10	.20	+	SS 10	FF 10	.20	+	SS 13	FF 7	1.	+	SS 19	FF 1	.02	−	SS 16	FF 4	.30	−
III	SF	SF 17	FS 3	.001	+	SF 16	FS 4	.001	+	SF 12	FS 7	.05	+	SF 18	FS 2	.001	+	SF 16	FS 4	.001	+
	FS	SF 4	FS 16	.001	+	SF 8	FS 12	.05	+	SF 5	FS 15	.001	+	SF 11	FS 9	.50	+	SF 10	FS 10	.20	+
IV	SS	SF 4	FS 11	FF 3	.05	−	SF 4	FS 10	FF 3	.10	−	SF 6	FS 8	FF 4	.70	?	SF 12	FS 6	FF 0	.01	−	SF 13	FS 5	FF 0	.001	−
	SF	SS 6	FS 1	FF 11	.02	−	SS 8	FS 2	FF 7	.20	+	SS 5	FS 3	FF 9	.20	?	SS 17	FS 0	FF 0	.001	−	SS 15	FS 0	FF 2	.001	−
	FS	SS 17	SF 0	FF 1	.001	+	SS 12	SF 3	FF 3	.02	+	SS 7	SF 5	FF 6	.90	?	SS 13	SF 5	FF 0	.001	+	SS 13	SF 3	FF 2	.01	+
	FF	SS 5	SF 10	FS 2	.10	−	SS 4	SF 13	FS 1	.01	−	SS 7	SF 9	FS 2	.20	+	SS 10	SF 4	FS 4	.20	−	SS 7	SF 5	FS 6	.90	?

* Yates' correction has been applied in every case where an expected frequency was less than ten and $df=1$.

** The signs indicate whether or not the direction of choice was as originally predicted on the basis of all choices being real. + indicates that it was; − indicates that it was not.

limitation upon the effectiveness of success and failure in the formation of groups is in terms of the free movement, both physical and psychological, permitted the individual in the society or culture. In a setting where geographical mobility is not grossly restricted, new interpersonal contacts would continually occur providing the members of such a society made use of the permitted mobility. This condition is, itself, not sufficient to satisfy the requirements. There must also be psychological mobility. In any society which severely limits the possible interpersonal relationships of its members by caste, clan, or any rigid autocratic means, it would not matter how much physical mobility existed, for the members of such societies would not be free to operate in new groups, i.e., make new grouping choices. Thus, before success and failure may be said to be effective, physical and psychological mobility must coexist: the more mobility, the greater the effectiveness of success and failure as an independent variable because there would be continually occurring new interpersonal actions.

Another factor of mobility must also be discussed. Beyond choice mobility, the interchange of communications dealing with new social ideals, empirical knowledge of social events and physical phenomena with their associated technological changes means a continual shift in the needs of its members and, consequently, their success-failure definitions. It is suggested that the more rapidly such a change occurs, the more effective success and failure will be from moment to moment.

The above statements related to the applicability of the results may be summarized as follows: the degree of psychological free movement (including ease of communication) and the number of individuals involved furnish the dimensions of the limiting cases of the effectiveness of success and failure in group formation. It is suggested, however, that these results have some measure of validity in grosser sociological situations such as those of Yankee City, Cornerville, etc. The arguments for this position and the limitations imposed upon it are given below.

It is assumed that in collections of individuals the amount of information available to and capable of being assimilated by a given individual about all other individuals becomes less as the size of the collection increases. Additionally, the number of interpersonal relationships which can be established in a unit of time has an upper limit. Above this upper limit, it is argued that positions on a commonly defined success-failure continuum come to be assigned not to individuals but to sets of individuals in a vague sort of statistical concept (e.g., classes and stereotypes such as "Riverbrookers," "Okies," etc.). Thus, in this statistical sense, the success-failure continuum continues to operate and expectations of task success are established in terms of assignments of individuals to classes upon the over-all continuum.

Over and above this specific issue, however, the lack of communication between individuals in large collections makes the over-all success-failure definition less likely to be common to all the members of the collection and this lack of commonality would increase as the amount of communication decreased. This factor would further limit the effectiveness of success and failure as an operative variable in large collections.

CONCLUSIONS

General Conclusions

1. The formulations, "level of aspiration" and "expectation," permit the prediction of individuals' grouping choices when a specific success-failure continuum is operative. Hence, they also permit prediction of the probable emergence of groups.

2. It would appear that two additional postulations beyond those used in this investigation will make such predictions more efficient. These are:

(a) that task success needs are, in the initial stages of repetitive success and/or failure, greater than choice success needs and

(b) that the first establishment of expectations of success and failure with other individuals via labeling as individuals has a greater effect upon the strength of such expectations than do subsequent single labelings after the formation of work units.

Specific Conclusions

1. When a random collection of individuals is distributed along a success-failure continuum so that some succeed and some fail as individuals, both the successes and the failures tend to group around the successes if group

action is required for further activity in the same dimension as that in which the success-failure continuum was previously operative.

2. If the typically reciprocal choices of the successes are made to operate to the exclusion of the failures and the successes succeed when working together, then

(a) the repeated successes tend to continue their grouping around other repeated successes when continued group activity along the same line of endeavor is required, whereas

(b) the repeated failures tend to distribute their choices between their fellow repeated failures and the repeated successes.

3. If the typically reciprocal choices of the successes are made to operate to the exclusion of the failures and the successes fail when working together while the failures succeed when working together, then

(a) the success-failures tend to continue their grouping around other success-failures, whereas

(b) the now failure-successes tend to form groupings around other failure-successes.

4. If the typical choices of successes by failures are made to operate (thereby blocking the typically reciprocal choices of the successes) and one of the mixed work units of successes and failures succeeds when working together while the other fails while working together, then

(a) the success-successes tend to continue their choices of the initial successes although these individuals are now success-failures, whereas

(b) the success-failures tend to continue their choices of the now successes-successes, (these choices do not attain statistical significance), whereas

(c) the failure-successes choose success-successes, whereas

(d) the failure-failures tend to distribute their choices between the success-successes and the success-failures (these choices do not attain statistical significance).

5. The grouping tendencies evidenced in 1 through 4 above tend to generalize to other activities than that of task dimension. It is suggested that such generalization is limited by the value placed upon these other activities in other frames of reference.

REFERENCES

1. FRENCH, J. R. P. Organized and unorganized groups under fear and frustration. *Univ. Ia. Stud. Child Welf.*, 1944, **20**, Pt. 5.
2. GOLDSCHMIDT, W. *As you sow.* New York: Harcourt, Brace, 1947.
3. HANFMANN, EUGENIA, & KASANIN, J. *Conceptual thinking in schizophrenia.* Washington: Nervous and Mental Disease Publishing Co., 1942.
4. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
5. KLEIN, G. S., & SCHOENFELD, N. The influence of ego-involvement on confidence. *J. abnorm. soc. Psychol.*, 1941, **36**, 249–258.
6. LEWIN, K., DEMBO, TAMARA, FESTINGER, L., & SEARS, PAULINE S. Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald Press, 1944.
7. LEWIS, D., & BURKE, O. J. The use and misuse of the chi-square test. *Psychol. Bull.*, 1949, **46**, 433–489.
8. LISSNER, K. Die Entspannung von Bedürfnissen durch Ersatzhandlungen. *Psychol. Forsch.*, 1933, **18**, 218–250.
9. LOOMIS, S. D., & GREEN, A. N. The pattern of mental conflict in a typical state university. *J. abnorm. soc. Psychol.*, 1947, **42**, 342–355.
10. LYND, R. S., & LYND, HELEN M. *Middletown.* New York: Harcourt, Brace, 1929.
11. LYND, R. S., & LYND, HELEN M. *Middletown in transition.* New York: Harcourt, Brace, 1937.
12. RASHKIS, H., CUSHMAN, J. F., & LANDIS, C. A new method for studying disorders of conceptual thinking. *J. abnorm. soc. Psychol.*, 1946, **41**, 70–74.
13. SNEDECOR, G. S. *Statistical methods.* Ames, Iowa: Iowa State College Press, 1940.
14. TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Century, 1932.
15. WALLER, M. The rating and dating complex. *Amer. sociol. Rev.*, 1937, **2**, 727–734.
16. WARNER, W. L., & LOW, J. O. *The social system of the modern factory.* New Haven: Yale Univ. Press, 1947.
17. WARNER, W. L., & LUNT, P. S. *The social life of a modern community.* New Haven: Yale Univ. Press, 1941.
18. WARNER, W. L., & LUNT, P. S. *The status system of a modern community.* New Haven: Yale Univ. Press, 1942.
19. WARNER, W. L., MEEKER, M., & EELLS, K. *Social class in America.* Chicago: Science Research Associates, 1949.
20. WARNER, W. L., & SROLE, L. *The social systems of American ethnic groups.* New Haven: Yale Univ. Press, 1945.
21. WHYTE, W. F. *Street corner society.* Chicago: Univ. of Chicago Press, 1943.
22. WRIGHT, M. E. The influence of frustration upon the social relationships of young children. *Psychol. Bull.*, 1941, **38**, 710. (Abstract.)

Received February 26, 1951.

MUSIC AND THE GENERAL ACTIVITY OF APATHETIC SCHIZOPHRENICS

CLYDE G. SKELLY AND GEORGE M. HASLERUD [1]
University of New Hampshire

ATTEMPTS to modify the activity of human beings by music have interested psychologists for many years. Tarchanoff (4) had early noted that "music exercises an influence on the muscular activity of man, increasing or decreasing it according to the nature of the melodic stimuli." On the introspective side, Weld (13) found that whatever visual imagery is present while listening to music is always imagery of movement. The earlier studies, however, did not lead to any consistent conclusion, based as they were on small samples, usually only a single case, little or no control of conditions, and general impressions rather than actual experimental evidence.

The last quarter century has brought more adequate experimentation on the motor side of the psychology of music, but these studies have largely involved the speeding up of such voluntary activities as handwriting and typing (5), delaying fatigue and increasing mental efficiency in industry (3), concentration and output in an architectural drafting room (6). Phares (11) showed that both gay and sad music markedly affected the psychogalvanic reflex. However, not until Altshuler (1) reported on two groups of four psychotic patients receiving hydrotherapy was music applied clinically in any sort of experimental setup. When programs arranged according to his iso-theory were played daily by an actual musician behind a screen, as much reduction in activity occurred with music as with hydrotherapy, and about 35 per cent reduction was observed as long as 15 minutes after the music ceased. He concluded, "Music seems a useful agent in decreasing output of disturbed and inaccessible mental patients both alone and in conjunction with various forms of hydrotherapy." Ogden (10), however, reviewing the entire field as portrayed in Schullian and Schoen's *Music and Medicine*, wrote, "Unhappily the content of the essays does little to demonstrate the value of music as a therapeutic agent or even to state 'the problems and institute a scientific approach to their solution.'"

[1] Adapted from a master's thesis written by C. G. S., supervised by G. M. H., on file in the University of New Hampshire library; the revision a joint product.

THE EXPERIMENT

Purpose

The present experiment attempts to meet Ogden's criticism by securing the essentials of a scientific setup that can be duplicated anywhere:

1. Stimulation by music of objectively determined mood value with programs arranged according to a definite plan.
2. Adequate method of quantitative recording of level of activity.
3. Cancelling out of individual differences by making each subject his own control.
4. Sufficient stability of level of activity when without music to detect changes with music.
5. Sufficient numbers to use a statistical test of significance.

Altshuler's data are inadequate to support his conclusion about the value of music in therapy because of neglect of points 4 and 5. Moreover, to insure reproducibility of results, both the stimuli and the recording of activity need to be much more systematically defined than we find in Altshuler.

Subjects

Thirty-nine female mental patients, 19 to 56 years of age, who had been diagnosed as schizophrenics of apathetic mood (blunted emotionality, lack of feeling, and impassivity) were used in the present experiment. All had been admitted to the hospital within the past year and showed no signs of intellectual deterioration. Their unchanging, chronic inactivity seemed promising for point 4.

Stimuli

In order to remove the bias of music selected only by the writers, a group of 81 recorded musical selections which were fairly constant in melody, rhythm, and tempo were selected, ranging from classical to near-popular. They were played to various groups of a total of 402 college students untrained in music, each piece being rated by at least 60 subjects on a scale based on the work of Rigg (12), Hevner (7), and Campbell (2). These workers had found that college students could rate the general mood of a musical composition, that is, whether it is sad, gay, happy, lively, or the like.

Hevner stresses the importance of rhythm and harmony in such ratings.

The rating scale consists of seven points with the following definitions: 1—stimulating, stirring, inciting, impelling, uplifting, elating, animated, tripping; 4—relaxing, restful, reposeful, tender, peaceful, gentle, sleepy, mellow; 7—depressing, oppressing, blue, dejecting, dismal, despondent, dispirited, despairing. Points 2, 3, 5, and 6 were left undefined but understood as moods proportional between the defined ones.

After all 81 selections had been rated by the groups of students, the median and semi-interquartile range were calculated and those with a Q of more than .75 discarded. The remaining 39 instrumental selections shown in Table 1 were then arranged into programs according to the iso-theory of Altshuler (1). This requires starting with music like the patient's mood and then gradually changing to a different one, in this experiment a gradual change from less active with high medians to more active with low medians. A sample program consisted of selections 38, 37, 31, 15, 6, 3, 13, 24, 37. These programs, lasting twenty to thirty minutes, came through the room's regular radio loud-speaker with the records played inconspicuously in an anteroom.

TABLE 1

PIECES USED IN PROGRAMS DURING MUSIC SESSIONS

No.	SELECTION	COMPOSER	SCALE VALUE	Q
1.	Tiger Polka	Uryga-Sandel	1.39	.59
2.	Perpetuum Mobile	Strauss	1.50	.56
3.	Helena Polka	Wladyslaw	1.55	.27
4.	Polonaise In A Major	Chopin	1.58	.73
5.	Jolly Coppersmith	Peter	1.60	.30
6.	Country Gardens	Grainger	1.62	.65
7.	Hail To the Spirit of Liberty	Sousa	1.62	.31
8.	Young Widow	Svec	1.64	.32
9.	Bumble Bee, The	Rimsky-Korsakow	1.78	.49
10.	Fiddle Faddle	Anderson	1.80	.50
11.	Etude No. 5 in G-Flat Major	Chopin	1.89	.66
12.	Second Movement, Organ Concerto No. 10 in D Minor—Allegro	Handel	1.94	.60
13.	Jazz Legato	Anderson	2.22	.62
14.	Zapateado	Sarasate	2.32	.51
15.	Ever Or Never	Waldteufel	2.51	.43
16.	Skater's Waltz, The	Waldteufel	2.68	.56
17.	Waltz In A-Flat Major	Chopin	2.79	.61
18.	Pizzicato Polka	Strauss	2.80	.43
19.	Wine, Women, and Song	Strauss	2.90	.57
20.	Southern Roses Waltz	Strauss	3.04	.65
21.	Voices of Spring	Strauss	3.12	.65
22.	Seamen's Waltz	Hellstrom	3.33	.65
23.	Missouri Waltz	Logan-Shannon	3.53	.72
24.	Gold and Silver Waltz	Lehar	3.60	.60
25.	Beautiful Ohio	Earl-MacDonald	3.64	.69
26.	Minuet In G	Beethoven	3.67	.61
27.	La Paloma	Yradier	3.69	.74
28.	Viener Blut	Strauss	3.69	.61
29.	Humoresque	Dvorak	3.71	.70
30.	Hear My Song Violetta	Bernier-Luckesch	3.82	.68
31.	First Love	Lehar	3.90	.63
32.	Trailing Arbutus, The	Friedman	4.07	.48
33.	Für Elise	Beethoven	4.56	.48
34.	Waltz In C-Sharp Minor	Chopin	4.69	.73
35.	Liebestraum	Liszt	4.75	.63
36.	Cradle Song	Brahms	4.82	.57
37.	On Wings Of Song	Mendelssohn	4.97	.62
38.	Barcarolle-Tales of Hoffmann	Offenbach	5.33	.65
39.	Sea Murmurs	Castelnuovo-Tedesco-Heifetz	5.52	.75

Recording

A quantitative measure of the general activity of *Ss* was obtained by time sampling, using a 7-point rating scale based on Newman (9) with these definitions: 1—overtly active practically all the time, extensive gross bodily movements, aggressive contacts with the physical environment, eager, animated bodily movements; 4—activity shown by *Ss* which one would expect to find in an audience listening to music, picking up a magazine, moving feet back and forth, nodding head to music, rocking slowly back and forth in a chair, fixing clothes, turning

pages of a magazine, wiggling back and forth in chair; 7—no overt activity, stationary, attitude of indifference, apathetic, no bodily movements. The intermediate points were left undefined as in the scale for selection of music. With each *S* observed for ten seconds on each rating, the group was sampled several times each session. At the 10:00 A.M. period the time samplings were made at the middle of the program when the livelier music was playing.

Controls on Recording

In order not to be influenced by them, previous ratings were kept covered. After the rating blank was completed, it was checked and filed by someone other than the rater and was not consulted until the end of the experiment.

To check the reliability of the ratings made by the experimenter, another observer (another psychologist) rated *Ss* for fifty samplings at the same time as the experimenter, both synchronizing their watches. The correlations between these ratings range between .70 and .88 for the various sessions.

Procedure

All samplings were made in an occupational therapy room to which the patients were accustomed. Before ratings were attempted, the experimenter became familiar with *Ss* and the *Ss* with the experimenter for a period of 12 days. When *Ss* came to the occupational therapy room, they were allowed to do as they wished—sit in a chair, look at books, or sew. The time samplings were not started until *Ss* had settled to their work. Because these patients were apathetic, many of them just sat in their chairs, looked at magazines—much the same activity that they would have shown on the ward. If for any reason an *S* did not want to come to the occupational therapy room, she was not forced to do so, and the sampling of the activity of that *S* was dropped for that day.

Sampling ratings were made five times a week at 8:00 A.M., 10:00 A.M., and 4:00 P.M. In order to determine a base level of activity, before the main experiment began a total of 117 ratings were made at each of the three times on each *S* over a two-week period in the room where music would later be given. Then, for over a month a total of over one hundred experimental ratings were made at each time of day on each *S*. Music was played at the 10:00 A.M. period, with the Before-Music (8:00 A.M.) and the After-Music (4:00 P.M.) ratings to indicate changes and duration.

RESULTS

The group results are shown in Tables 2 and 3. Each *S* had six scores, each the mean of 99 to 117 samplings of general activity, which were the scores of base level at 8:00 A.M., 10:00 A.M., and 4:00 P.M. (preliminary No-Music period), and the experimental series 8:00 A.M. (Before-Music), 10:00 A.M. (With-Music), and 4:00 P.M. (After-Music). Thus, the means in Tables 2 and 3 are actually the central tendency of 4089 to 4563 time samplings.

TABLE 2

DIURNAL SHIFTS IN ACTIVITY OF 39 APATHETIC PATIENTS

ACTIVITY	TIME OF SAMPLING	MEAN	SD	*r*	*t*
Base Level compared with Base Level	8:00 A.M.	5.12	.73	.95	1.29
	10:00 A.M.	5.07	.79		
Base Level compared with Base Level	10:00 A.M.	5.07	.79	.94	2.03
	4:00 P.M.	5.16	.69		
Base Level compared with During-Music	10:00 A.M.	5.07	.79	.88	10.63
	10:00 A.M.	4.41	.80		

TABLE 3

AMOUNT AND DURATION OF EFFECT OF MUSIC ON ACTIVITY OF 39 APATHETIC PATIENTS

ACTIVITY DURING PERIOD SAMPLED	M.	SD	*r*	*t*
Base Level with Before-Music	5.13	.72	.93	.18
	5.14	.89		
Before-Music with During-Music	5.14	.89	.96	17.86
	4.41	.80		
During-Music with After-Music	4.41	.80	.95	15.38
	5.10	.88		
Before-Music with After-Music	5.14	.89	.98	1.48
	5.10	.88		

The activity scores range from 4.29 to 6.44 for the various *Ss*, and the median standard deviation is .59 on the 7-point scale. Each *S* was remarkably stable in apathetic behavior during the several months of this study and even showed little diurnal change, as attested by the *r*'s of .94 and .95 in Table 2. The only shift was a slight but insignificant increase in activity at 10:00 A.M. ($t = 1.29$), which had been chosen as the time for playing music in the experimental series. Since, however, each *S* was her own control, the very great and significant change at 10 o'clock when music was played, shown by a *t* of 10.63 in the last column of Table 2, clearly differentiates the base level sampling from the experimental series sampling.

Table 3 indicates that only while music was

playing was there greater activity. When the After-Music and Before-Music differences are compared, there is an insignificant *t*. In contrast, when the During-Music and After-Music, and the Before-Music and During-Music time sampling scores are compared, there are highly significant *t*'s.

To check the findings from the group data an analysis of the individual records of each of the 39 *Ss* was undertaken. For each patient, the differences in pairs of ratings, e.g., the 10:00 A.M. minus the 8:00 A.M., were tabulated and the mean and standard deviation of the differences determined and interpreted for correlated samples according to McNemar (8). Meeting a *t* of 3.0 as the required level of significance beyond the 1 per cent point, all *Ss* showed significant increase of activity when music of such variety and in such programs as herein described was played to them. Further, a study of the individual records indicates that 35 *Ss* regularly went back to their previous base level, and six hours later (4:00 P.M.) the effect of the music was indistinguishable. Only four *Ss* had a *t* of greater than 3.0 when the activity at 8:00 A.M. was compared with that at 4:00 P.M. In two there was a slight persistence of activity, but this was counterbalanced by two who went below morning level; because of the high *r*, these differences were significant. For the majority, both the individual and group data indicate that music makes a difference in activity only while it is being played or shortly thereafter.

DISCUSSION

That the changes in activity during music were not due to diurnal rhythm nor to a conditioning effect of the experimenter's presence is demonstrated by the relatively narrow variability in all the nonmusic sessions, for all three times of day on the control base-level samplings and for the 8:00 A.M. and 4:00 P.M. periods of the experimental series.

Altshuler (1) found that music could decrease the activity of eight disturbed schizophrenic patients, and the present experiment shows that music significantly increased the general activity of 39 apathetic schizophrenics. This change in activity level of mental patients might give hope that music could be a form of psychotherapy. However, our results would cast doubt on such value of music by itself. When there is an insignificant carry-over of the effects of music for six hours and a continuing stable morning level of activity over several months during our experiment, one must conclude that music can probably be only an adjuvant to psychotherapy. For example, programs of the kind used in this experiment could be employed to regulate the activity of patients while in group psychotherapy, or in the dining halls of mental hospitals to increase the eating activity of retarded mental patients, or as a technique in occupational therapy to get patients to participate.

The music in this experiment was classical to near-popular. It should be determined whether modern popular music of the "Hit Parade" variety could give a more persistent effect with its simpler repetitive melodies. A check should also be made with excited patients to see whether the effect of quieting music would last the six hours from 10:00 A.M. to 4:00 P.M.

SUMMARY AND CONCLUSIONS

The purpose of this experiment was to study the effect of music, ranging from classical to near-popular, on the general activity of apathetic schizophrenics.

Music that had been calibrated on a 7-point scale for mood by 402 normal college students was played in programs arranged according to the iso-theory.

A 7-point rating scale was used to determine the general activity of patients with and without music. A check of their base level of general activity without music showed it, through high *r*'s, to be very stable.

1. Both individually and as a group the patients showed significant increase in activity when livelier music was played to them.

2. For 35 of the 39 patients, there was no carry-over effect of the music after a period of six hours; two of the 39 patients showed slight persistence of activity and two a decrease slightly below morning level.

3. It is concluded that the therapeutic effects of music such as used in this experiment are temporary, and its probable main use would be as an adjuvant to other therapy.

REFERENCES

1. ALTSHULER, I. M. A psychiatrist's experience with music as a therapeutic agent. In D. M. Schullian, and M. Schoen (Eds.), *Music and medicine.* New York: Henry Schuman, 1948.
2. CAMPBELL, J. G. Basal emotional patterns experienced in music. *Amer. J. Psychol.*, 1942, **55**, 1–17.
3. CLARK, K. *Music in industry.* New York: National Bureau for the Advancement of Music, 1929.
4. DISERENS, C. M. Reactions to musical stimuli. *Psychol. Bull.*, 1923, **20**, 173–199.
5. DISERENS, C. M. *The influence of music on behavior.* Princeton: Princeton Univ. Press, 1926.
6. GATEWOOD, E. L. An experiment in the use of music in an architectual drafting room. *J. appl. Psychol.*, 1921, **5**, 350–358.
7. HEVNER, K. Experimental study of the elements of expression in music. *Amer. J. Psychol.*, 1936, **48**, 246–268.
8. MCNEMAR, Q. *Psychological statistics.* New York: Wiley, 1949, 73–75.
9. NEWMAN, F. B. The adolescent in social groups; studies in the observation of personality. *Appl. Psychol. Monogr.*, 1946, 9.
10. OGDEN, R. N. Review of D. M. Schullian, and M. Schoen (Eds.), *Music and medicine. Amer. J. Psychol.*, 1934, **61**, 595.
11. PHARES, M. L. An analysis of musical appreciation by means of psychogalvanic reflex techniques. *J. exp. Psychol.*, 1934, **17**, 119–140.
12. RIGG, M. The expression of meanings and emotions in music. *Psychol. Bull.*, 1940, **37**, 556.
13. WELD, H. P. An experimental study of musical enjoyment. *Amer. J. Psychol.*, 1912, **23**, 245–308.

Received March 8, 1951.

INDEX OF NAILBITING IN CHILDREN [1]

ANTHONY J. MALONE AND MAURY MASSLER [2]
University of Illinois

A REVIEW of the literature led to the conclusion that nailbiting, like thumbsucking, might be (*a*) a normal action if indulged in occasionally and one which merits little attention by the clinician; (*b*) a simple tension releasing mechanism if the child is addicted to it to a mild degree; or (*c*) a sign of severe internal tension if the action is severe and persistent (4, 5). During that review it became apparent that very little, if any, attention had been given to the significance of different degrees of nailbiting. No differentiation was made between mild, occasional nailbiting and severe, persistent nailbiting. It seemed desirable, therefore, to restudy the question, keeping in mind the distinction between different degrees of nailbiting.

The purpose of this investigation was (*a*) to describe a simple, quantitative method of assessing different degrees of nailbiting and (*b*) to accumulate statistically reliable data on the prevalence of different degrees of nailbiting in children.

METHOD

Subjects

A total of 4587 children, five to 18 years of age, were examined. The subjects were of both sexes from Chicago and Chicago suburban schools living in the middle socioeconomic levels.

Methods of Assessment

Severity of nailbiting. The condition of each finger was recorded according to the scale of severity described in Table 1 and illustrated in Fig. 1.

Index of nailbiting. A total score or "index" was derived for each person by adding the numerical rating for each finger of both hands. This procedure gave each child a nailbiting "score" between 0 and 30.

Each person biting the nails was asked only two simple questions: (*a*) When do you bite your nails? (*b*) Why do you bite your nails?

[1] From the Division of Graduate Pedodontics of the College of Dentistry.

[2] Professor of Graduate Pedodontics.

RESULTS

Number of Persons Biting Their Fingernails

Prevalence. Forty-one per cent of the 4587 children (ages five to 18 years) bit one or more of their fingernails (Table 2). This figure is somewhat higher than previously reported by Wechsler (1931) (Fig. 7).

Age incidence. According to Wechsler (1931) and Billig (1941), very few children bite their nails before the age of five years. This investigation therefore began with children five years of age and over. It was discovered that a considerable number (38 per cent) of the five-year-old children in this group bit their nails.

The prevalence of nailbiting (one to 10 fingers) rose sharply from 38 per cent at five years to 61.6 per cent at eight years. From eight to 10 years of age, the number of children who bit their fingernails remained at a relatively flat level (about 60 per cent). After 10 years of age, there was a regular decrease in the number of children who bit their fingernails (from 60 per cent at 10 years to 28 per cent at 18 years, see Table 2 and Fig. 2).

Sex differences. From five to 10 years of age, there did not appear to be any significant differences between males and females. The respective curves were rather irregular and overlapped frequently (Fig. 3). Since the irregularity of the individual curves might mask a real sex difference, the data were subjected to statistical analysis. It was found that there were, statistically, no significant differences in the number of boys and girls biting their nails between the ages of five to 10 years ($p=.07$).

From 10 to 18 years of age, however, the curves showing the number of boys and girls biting one to ten fingernails at different ages appeared to be significantly different. The number of girls who bite their nails was regularly about 10 per cent less than the number of boys. The difference between

TABLE 1

A SCALE FOR MEASURING THE SEVERITY OF FINGERNAIL BITING

DEGREE OF FINGERNAIL BITING	DESCRIPTION
0 (not bitten)	Free margin intact.
1+ (mildly bitten)	Free edge of nail irregular but reasonably intact. Fingernail biting confirmed on questioning.*
2+ (moderately bitten)	Free margin of nail absent. Regular biting confirmed on questioning.
3+ (severely bitten)	Fingernail bitten beyond the free edge; nail margin below the soft tissue border (Fig. 1).

*In cases of doubt, the fingernail was assumed to have been accidentally broken and was not scored as bitten. This seldom occurred.

since the condition of the fingernail will reflect the biting for a considerable length of time.

Twenty-eight per cent of the children (or 68.0 per cent of those who bit their nails) bit all 10 fingernails. Only a very small number (0.55 per cent to 1.8 per cent) bit 1 to 9 fingernails (Table 3 and Fig. 4). This grouping indicated clearly that when a child bit his nails, the tendency was (in two out of three cases) to bite all ten fingernails.

The bimodal distribution, as shown in Fig. 4, suggests that there is a "natural" grouping of the children into three categories (see Table 4). Those children who bit none of their fingernails are obviously "non-nailbiters." Children who bite all 10 fingernails may be called "definite" nailbiters. Children who bite 1 to 9 fingernails are either transitional

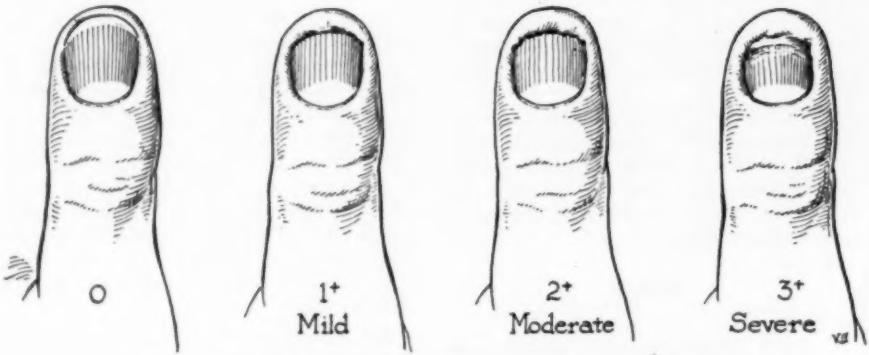

FIG. 1. DIAGRAM SHOWING DIFFERENT DEGREES OF FINGERNAIL BITING

the means for boys and girls was analyzed statistically and found to be significant (p=.03).

The curve showing the number of girls who bite their nails begins to decelerate sharply at the age of nine years. The curve for boys does not begin to decelerate until the age of 11 years. This probably reflects the earlier social consciousness of girls (2).

Number of Fingernails Bitten

The majority of the children (59 per cent) did not bite any of their fingernails as judged from the condition of the nails at the time of the examination. It is unlikely that these children bit their nails even occasionally between the other two categories or are an indeterminate group and may be termed "indefinite" nailbiters.

Age trends. The number of non-nailbiters, definite (10) nailbiters and indefinite (1 to 9) nailbiters were graphed according to age (Fig. 2 and Fig. 3).

The number of definite (10) nailbiters increased from a low of 28 per cent at 5 years of age to a peak of 50 per cent at the age of 10 years, decelerating rapidly thereafter to 15 per cent at 18 years of age (Fig. 2). This reaffirmed the trend shown by the percentage of all (1 to 10) nailbiting children. However, by refining the data into "definite" (10) and "indefinite" (1–9) fingernail biters, the

TABLE 2

PREVALENCE OF NAILBITING IN 4587 CHICAGO SCHOOL STUDENTS (AGES 5 TO 18 YEARS)

AGE IN YEARS	COMBINED		BOYS		GIRLS	
	NUMBER EXAMINED	% BITING	NUMBER EXAMINED	% BITING	NUMBER EXAMINED	% BITING
5	168	38.0	81	42.0	87	34.4
6	292	45.2	155	40.5	137	50.5
7	189	45.4	84	41.6	105	48.5
8	198	61.6	105	63.7	93	59.2
9	150	60.0	75	57.5	75	62.7
10	188	60.6	88	60.2	100	61.0
11	156	57.4	78	64.1	78	51.5
12	166	51.2	88	58.0	78	43.6
13	173	49.1	84	56.0	89	42.6
14	629	42.4	219	53.8	410	36.6
15	760	38.9	317	44.4	443	35.1
16	694	32.4	341	37.8	353	27.1
17	689	28.0	335	38.2	354	18.3
18	135	28.1	80	31.3	55	23.7
Total	4587	41.1	2130	46.1	2457	36.7

age trends in nailbiting became more regular (Fig. 3). The peak of definite (10) nailbiting lies at 10 years of age instead of somewhere between eight and 10 years as suggested by the data which group together both the definite and indefinite nailbiters (Fig. 2 and Fig. 3).

The curve representing the number of indefinite (1–9) fingernail biters is very irregular (Fig. 3). There is no significant trend with age—reaffirming the hypothesis that this is an "indefinite" group and one which is probably not significant to the total picture of fingernail biting.

Sex differences. The anlaysis of all children biting their fingernails suggested that a real sex difference existed after 10 years of age. Analysis of the definite (10) fingernail biters emphasizes this fact and shows that sex differences begin even earlier, at about

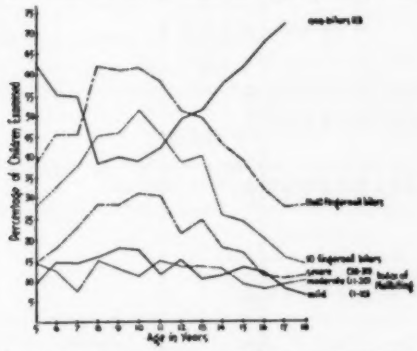

FIG. 2. GRAPH SHOWING AGE TRENDS IN FINGERNAIL BITING ACCORDING TO NUMBER OF FINGERNAILS BITTEN (0, 1—10, 10) AND SEVERITY OF NAILBITING (MILD, MODERATE AND SEVERE)

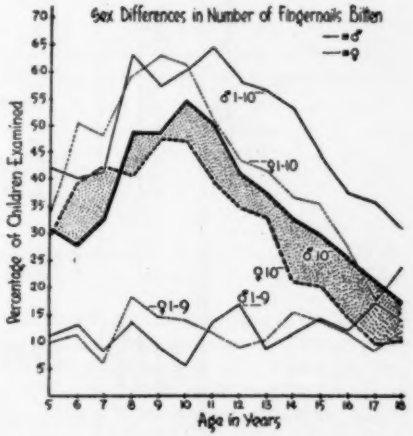

FIG. 3. SEX DIFFERENCES IN THE NUMBER OF FINGERNAILS BITTEN
(The numbers refer to the number of fingernails bitten.)

TABLE 3

DISTRIBUTION OF CHILDREN ACCORDING TO THE NUMBER OF FINGERNAILS BITTEN

AGE IN YEARS	TOTAL NO. OF CHILDREN EXAMINED	NUMBER OF CHILDREN BITING 0–10 FINGERNAILS												BOYS (%)	GIRLS (%)
		0		1	2	3	4	5	6	7	8	9	10	10	10
		No.	%										No. %		
5	168	104	(61.9)		1	3	4	3	2	1	2	1	47 (28.0)	(30.8)	(25.2)
6	292	160	(54.7)	3	4		6	6	5	2	5	4	97 (33.1)	(27.6)	(39.4)
7	189	103	(54.4)	1	1		3	2	4	1	2		72 (38.1)	(33.3)	(42.0)
8	198	76	(38.3)	1		3	9	7	5	2	2	3	90 (45.2)	(49.3)	(40.9)
9	150	60	(40.0)	1	1	2	2	1	2	2	4	6	69 (46.0)	(49.0)	(48.0)
10	188	74	(39.3)		4	2	2	1	6		2	2	95 (50.5)	(54.5)	(47.0)
11	156	66	(42.3)		5	2	2	1	3	2	3	2	70 (44.8)	(50.0)	(39.7)
12	166	81	(48.7)		4	2	3	4	1	4	2	3	63 (38.6)	(41.0)	(34.6)
13	173	88	(50.8)	3	3		2	2	1		4	1	69 (39.7)	(47.7)	(32.6)
14	629	362	(57.5)	3	5	11	17	16	16	15	11	14	159 (25.4)	(32.9)	(21.2)
15	760	464	(61.0)	2	8	14	12	13	23	11	17	11	185 (24.6)	(29.5)	(20.5)
16	694	469	(67.5)	3	9	7	10	15	8	14	11	8	140 (20.2)	(25.7)	(14.7)
17	689	496	(71.9)	5	14	6	10	8	8	12	11	13	106 (15.4)	(20.5)	(9.6)
18	135	97	(71.8)	1		3	2	2		3	5	2	20 (14.8)	(17.5)	(10.8)
Total	4587	2700	(58.9)	25	58	55	84	80	84	69	80	70	1282 (28.0)	(32.1)	(24.3)
				%=(.55)	(1.2)	(1.2)	(1.8)	(1.7)	(1.8)	(1.5)	(1.7)	(1.5)			
				|——————————13.1%——————————|											

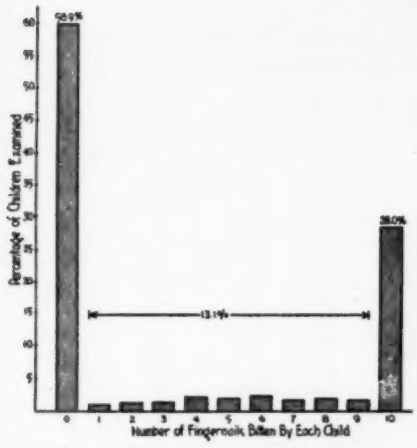

FIG. 4. DISTRIBUTION OF CHILDREN ACCORDING TO THE NUMBER OF FINGERNAILS BITTEN

eight years of age (Table 3 and Fig. 3). Girls begin to stop biting their fingernails at nine years of age, boys about one year later.

The indefinite (1–9) fingernail biters behave in an indefinite manner. The number of males and females biting 1 to 9 fingernails is very small. The trend is flat and very irregular. The curves overlap so that there appear to be no sex differences from five to 18 years. The irregularity of these curves again supports the thesis that the indefinite (1–9) fingernail biters are not significant in composing a true picture of fingernail biting.

Degree (Index) of Nailbiting

Each finger was assessed as to the degree (severity) of nailbiting: 0, 1+, 2+, or 3+.

TABLE 4

CATEGORIES OF CHILDREN FROM THE STANDPOINT OF MEASURES OF NAILBITING

	PERCENTAGES	
	TOTAL GROUP	NAILBITING GROUP
1. Non-nailbiters	59	
2. Children who bite all 10 fingernails	28	68
3. Children who bite 1 to 9 fingernails	13	32
	100	100

The sum (index) for each child was derived from the rating for each finger. Each child was thus rated according to the severity of nailbiting from 0 to 30.

Distribution of mild, moderate, and severe nailbiters. The distribution of the number of children with different degrees of nailbiting is shown in Table 5 and Fig. 5. There is a strong quadrimodal distribution.

TABLE 5

DISTRIBUTION OF CHILDREN ACCORDING TO INDEX OF NAILBITING

INDEX OF NAILBITING	NO. CHILDREN	% BITING	% MALES	% FEMALES
0	2700	59.0	53.8	63.2
1	14	0.3		
2	22	0.4		
3	34	0.7		
4	46	1.0		
5	38	0.8		
6	53	1.1		
7	38	0.8		
8	57	1.2		
9	42	0.9		
10	160	3.4	4.1	2.9
1–10		10.9		
11	29	0.6		
12	40	0.8		
13	35	0.7		
14	22	0.4		
15	32	0.6		
16	37	0.6		
17	16	0.3		
18	34	0.7		
19	27	0.5		
20	287	6.3	6.9	5.6
11–20		12.2		
21	13	0.2		
22	17	0.3		
23	4	0.08		
24	10	0.2		
25	13	0.3		
26	9	0.2		
27	18	0.3		
28	10	0.2		
29	9	0.2		
30	719	15.6	17.7	13.8
21–30		17.9		

As already pointed out, most of the children (59 per cent) had an index of 0, i.e., did not bite their fingernails at all. Those children who did bite their nails usually bit all 10 fingernails at three degrees of severity; mildly (10×1), moderately (10×2) or severely (10×3). Very few children showed mixed degree of biting between these four modal groups (Table 5 and Fig. 5).

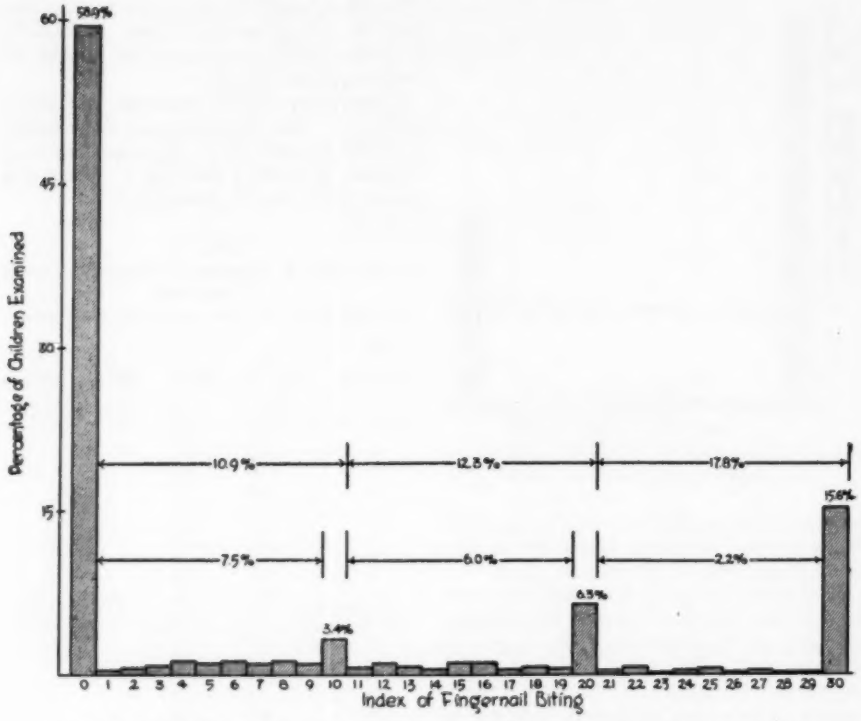

FIG. 5. DISTRIBUTION OF CHILDREN ACCORDING TO THE INDEX OF FINGERNAIL BITING (0–30)

Eleven per cent of the children bit with a severity from 1 to 10. Since 13 per cent of the children were indefinite (1–9 fingernails) biters, it was thought that this group is probably the same as those previously categorized as "indefinite nailbiters." Re-examination of the raw data substantiated this thesis: that the indefinite (1–9 fingernails) biters bit to an index from 1 to 10.

Twelve and two-tenths per cent of the children bit their nails with an index from 11 to 20 while 17.9 per cent bit from an index of 21 to 30. These children (30 per cent) probably correspond to the previous group of (28 per cent) "definite nailbiters." Re-examination of the raw data confirmed the thesis: that, with a very few exceptions, children who bit 10 fingernails bit at an index from 11 to 30.

The striking modal concentration about the indices of 20 and 30 is very significant since it emphasizes the fact that the definite nailbiters (children who bite 10 fingernails) tend with but very few exceptions to bite all fingers uniformly at 2+ or 3+ (more of them at 3+ than 2+).

In view of the striking modality, the data for the definite (10) fingernail biters were re-examined, and it was found that among these children, 10 per cent bit mildly (1+), 21 per cent bit moderately (2+) and 57 per cent bit severely (3+). These figures suggest that most (57 per cent) of the definite (10) nailbiters bit their nails severely (3+). Some (21 per cent) bit moderately (2+) but only a few (10 per cent) of the 10-fingernail biters bit mildly (1+).

We may, therefore, define fingernail biters not only as being definite (10 fingernails) or indefinite (1–9 fingernails), but we may also categorize them as being mild fingernail

TABLE 6

DISTRIBUTION OF CHILDREN (5–18 YEARS) ACCORDING TO INDEX OF NAILBITING (1–10, 11–20, 21–30)

AGE IN YEARS	TOTAL NO. EXAMINED	ZERO (0)		MILD (1–10)		MODERATE (11–20)		SEVERE (21–30)	
		No.	%	No.	%	No.	%	No.	%
5	168	104	61.9	24	14.2	16	9.5	24	14.2
6	292	160	54.7	36	12.3	43	14.7	53	18.1
7	189	103	54.4	14	17.4	28	14.8	44	23.2
8	198	76	38.3	33	16.6	32	16.1	57	28.7
9	150	60	40.0	20	13.3	27	18.0	43	28.6
10	188	74	39.3	21	11.1	33	17.5	60	31.9
11	156	66	42.3	23	14.7	19	11.5	48	30.7
12	166	81	48.7	23	13.8	26	15.6	36	21.6
13	173	88	50.8	24	13.8	18	10.4	43	24.8
14	629	362	57.5	82	13.0	72	11.4	113	17.9
15	760	464	61.0	71	9.3	100	13.1	125	16.4
16	694	469	67.5	54	7.7	82	11.8	89	12.8
17	689	496	71.9	65	9.4	56	8.1	72	10.4
18	135	97	71.8	14	10.3	9	6.6	15	11.1
Total	4587	2700	59.0	504	10.9	561	12.2	822	17.9

biters (degree 1 to 10), moderate (11 to 20) or severe (21 to 30) fingernail biters.

Further, 13 per cent of the children bite their nails in an indefinite manner (1–9 fingernails), 11 per cent to a mild degree (index 1–10). Twenty-eight per cent of the children bite all 10 fingernails; 12 per cent to a moderate degree (index 11–20) and 18 per cent severely (index 21–30) (Table 7).

Age trends of mild, moderate, and severe nailbiting. The number of children in each of the three groups was graphed according to age in order to determine the trends in nailbiting (Fig. 2). This graph shows that there is no significant increase or decrease from five to 18 years of age in the number of children who bite their fingernails mildly (index 1–10) or moderately (index 11–20). However, the number of children who bite severely (index 21–30) rises from a low of 15 per cent at five years of age to a peak of 33 per cent at 10 years, decreasing to 10 per cent at 18 years. In other words, the age trend in the severe nailbiting group exactly duplicates the trend of the total group (per cent of children biting) and also the trend of the definite (10) fingernail biters.

From the above, the conclusion might be drawn that only those children who bite 10 fingernails severely (index 21–30) are the "true" fingernail biters. This is the significant group which markedly influences the trend for the total group. While the group of severe nailbiters is only 43.7 per cent of the total number of nailbiters, it weights the data for the whole group heavily since the other two groups of mild and moderate nailbiters show very little or no age trends. The data for the total group, therefore, in reality, reflect the behavior of the severe nailbiters—and not the behavior of the mild and moderate nailbiters (Fig. 2).

Whether the mild (index 1–10) nailbiters

TABLE 7

DISTRIBUTION OF DIFFERENT TYPES OF NAILBITERS

PERCENTAGE OF TOTAL NUMBER CHILDREN

	NAILBITING	
	Number of Fingernails Bitten	Degree (Index) of Fingernail Biting
Non-Nailbiters 59%	None	(0)
Nailbiters 41%	Indefinite (1–9 fingers) 13.0%	Mild (1–10) 11%
	Definite (10 fingers) 28.0%	Moderate (11–20) 12.2%
		Severe (21–30) 17.9%

should be considered as "nailbiters" now seems doubtful. Further analysis must establish the significance of this group as well as the group of moderate nailbiters (index 11–20).[3]

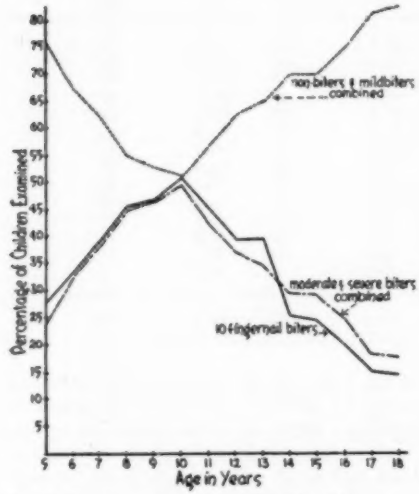

FIG. 6. COMBINATION OF MILD AND NON-NAILBITING, AND SEVERE AND MODERATE NAILBITING

Nailbiting and malocclusion. Nailbiting has often been indicated as an etiologic factor in malocclusion. The occlusion of the teeth of these children was, therefore, also carefully examined.

While malocclusion of the teeth and nailbiting were both found to be very common in this group, in no instance could the performance of the act be directly related to the displacement of a tooth or teeth.

DISCUSSION

Comparison with earlier studies. The overall frequency of nailbiting in this group of children was somewhat higher than that reported by Wechsler in 1931. The observational techniques used in both studies were approximately the same (although Wechsler did not record the severity of biting).

The number of children biting their nails at 5–6 years was larger in the present group (38 per cent) than in Wechsler's group (28 per cent). In addition, the peak of nailbiting in this group occurred at the prepubertal period (8–10 years of age) while Wechsler's group (Fig. 7) reached its peak at puberty

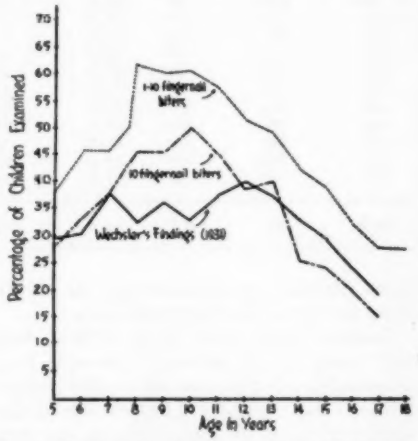

FIG. 7. COMPARISON OF PREVALENCE OF NAILBITING AS DETERMINED BY WECHSLER IN 1931 WITH PRESENT STUDY (1950)

(12–14 years of age). The present group of children, therefore, began to bite their fingernails earlier; more children bit their nails at every age period and the number of children biting their nails reached its peak earlier. These differences appear to be real and not accidental and might indicate that tensions in children have increased during the last two decades (1931 to 1950).

Definition of a nailbiter. Clinically, the nailbiter has been defined only as one who bites his nails. Only two types of children have been recognized in the past—those who bite their fingernails, and those who do not.

It is obvious that children who bite their nails do so to different degrees of severity. Examination and analysis showed that 41 per cent of the children bite their nails. Most of the children (2 out of 3) who bite their fingernails do so on all 10 fingers. Only a few (1 out of 3) bite irregularly from 1 to 9 fingernails. We may, therefore, designate

[3] In order to test whether it would be possible to identify the "true" nailbiter statistically, the data were rearranged in various ways. When the group composed of non-nailbiters and mild nailbiters is plotted against the group composed of moderate and severe nailbiters, the curves are almost mirror images. Note also the similarity between the group of 10-fingernail biters and the group composed of moderate and severe nailbiters (Fig. 6).

these two subgroups as "definite" and "indefinite" nailbiters.

If an attempt is made to score the degree of fingernail biting by the effects upon the nails (from slight to moderate to severe) it is possible to index (numerically) the severity of the fingernail biting. Such analysis showed that most children (18 out of 41 per cent) who bite their nails do so severely on all ten fingers (index 21–30). This group may be termed "severe" nailbiters.

Twelve out of 41 per cent bite to a moderate degree (index 11–20) while 11 out of 41.1 per cent bite mildly (index 1–10).

Analysis of the age trends of each one of the groups designated above indicated that the mild and moderate nailbiter did not influence the over-all trend in nailbiting and that *the severe nailbiter was the significant factor in the total picture of nailbiting.*

It is suggested, therefore, that only the child who bites all 10 fingernails severely be considered as the "true" fingernail biter. Further studies are indicated to more clearly define the significance of severe nailbiting as well as mild and moderate nailbiting. More exact definition may reveal more clearly the significance of the act.

Nailbiting in young adult males. A group of 120 male dental students were examined for nailbiting. Their ages ranged from 22 to 38 years with the mean at 25 years. It was found that 11 (9 per cent) of these young adult males bit their nails; 3 (2.3 per cent) bit mildly (10×1), 2 (0.8 per cent) bit moderately (10×2) and 5 (4.1 per cent) bit severely (10×3). One adult bit only two fingers in a peculiar fashion. These figures fit the trends shown in Fig. 2 and confirm the findings of Pennington (7).

Pattern of nailbiting. Bilaterality. When the data were analyzed for hand and finger preference in biting, it was discovered that nailbiting was bilaterally symmetrical: ninety-six and seven-tenths per cent of the nailbiters bit the fingers of each hand equally (showing no preference for the right or left hand). This finding probably indicates that nailbiting is an *unconscious* activity.

Finger preference. The order of preference in which fingers were bitten was (*a*) thumb, (*b*) index finger, (*c*) little finger, (*d*) middle finger, (*e*) ring finger.

The nailbiter usually began with the thumb, advanced to the index finger, skipped to the little finger, thence to the middle finger, and last to the ring finger. The reason for leaving the ring finger last in the order of biting may be explained by the fact that ligamentous attachments to the ring finger make its separation from the other fingers difficult, so that biting the nail of the third finger is more clumsy.

Reasons for nailbiting. Most of the children reported that they bit their nails during situations which made them nervous (watching exciting movies, during school examination, radio and television programs, family quarrels, and sport events). Twenty-five per cent of the nailbiters reported that they bit their nails during periods of boredom (enforced inactivity in school). Approximately 16 per cent of the nailbiters could give no reason for biting their nails.

SUMMARY AND CONCLUSIONS

An epidemiologic method for the assessment of nailbiting is presented. The method is simple, quick, and gives quantitative results which are reproducible and can be subjected to statistical treatment. This method revealed that the traditional method of assessing nailbiting (absence or presence) is not a good analytic approach. Evaluating a nailbiter according to the number of fingernails bitten and the degree of biting is more accurate.

Children may bite from 1 to 10 fingernails. The 10 fingernail biters are "definite" nailbiters; those who bite 1 to 9 fingernails constitute a smaller and an "indefinite" group. The degree of nailbiting may vary from mild to severe. A child may bite 10 fingers each mildly (1+), moderately (2+) or severely (3+). Most children who bite their nails bite all 10 fingers and to a severe degree.

Analysis of age trends showed that severe nailbiters were the significant factor in the total picture of trends in nailbiting. The mild and moderate nailbiters do not greatly influence the trend in nailbiting.

The *significance* of nailbiting is probably an individual problem and should be approached clinically. Future studies should be designed to distinguish the mild and moderate nailbiter from the severe nailbiter.

REFERENCES

1. BILLIG, A. L. Fingernail biting: The incipiency, incidence and amelioration. *Genet. Psychol. Monogr.*, 1941, **24**, 123–218.
2. COLEMAN, J. C., & MCCALLEY, J. E. Nailbiting among college students. *J. abnorm. soc. Psychol.*, 1948, **43**, 517–525.
3. MASSLER, M., & WOOD, A. W. S. Thumbsucking. *J. dent. Children*, 1949, **16**, 1–9.
4. MASSLER, M., & MALONE, A. J. Nailbiting—a review. *J. Pediat.*, 1950, **36**, 523–531.
5. MASSLER, M., & MALONE, A. J. Nailbiting—a review. *Amer. J. Orthod.*, 1950, **36**, 351–367.
6. OLSON, W. C. *The measurement of nervous habits in normal children*. Minneapolis: Univ. Minnesota Press, 1929.
7. PENNINGTON, L. A. Incidence of nailbiting among adults. *Amer. J. Psychiat.*, 1945, **102**, 241–244.
8. WECHSLER, D. The incidence and significance of fingernail biting in children. *Psychoanal. Rev.*, 1931, **18**, 201–209.

Received March 16, 1951.

FRUSTRATION, SATISFACTION, AND OTHER FACTORS AFFECTING THE ATTRACTIVENESS OF GOAL OBJECTS [1]

ROBERT J. FILER
Duke University

An important problem in the psychology of motivation is the determination of the conditions that affect the attractiveness of goal objects. Related to the conditions that influence the attractiveness of goal objects is the establishment of specific satisfiers and the establishment of individual preferences. In an attempt to account partly for the existence of individual preferences, Murphy posited a canalization process, "by which general motives (which are at first rather non-specifically related to a class of stimuli) tend upon repeated experience to become more easily satisfied through the action of the specific satisfier than of others of the same general class" (16, p. 162). Thus, consistent satisfaction of a need by a particular goal object leads to a preference for that object. Murphy's discussion of the canalization process emphasizes a need for research concerned with preferences for specific goal objects. In another context, Kurt Lewin, considering changes in needs and goals, stated, "Reaching a goal as well as not reaching it, may change the valences in a momentary or permanent way" (13, p. 823). He did not indicate the direction of the change nor did he further elaborate upon this specific proposition. The indefiniteness of his statement further emphasizes the importance of an experimental investigation dealing with the attractiveness of goal objects. Both Murphy's and Lewin's discussions raise important questions about the effects of the attainment and nonattainment of a specific goal object upon the object's attractiveness.

Although a survey of the psychological literature yields little experimental evidence that would permit one to state more specifically the effects of attainment and nonattainment of goal objects upon their attractiveness, there are several studies which seem to be related to the general problem of goal attractiveness. Some of these studies, which will be discussed in more detail below, have related goal preferences to the physical and temporal nearness of the goal. Other recent experiments have produced information about the attractiveness of activities and tasks. While the activity one does to obtain a goal object is distinct from his goal, it is possible that variables similar to those found important in the activity studies will influence the attractiveness of a goal object. Rosenzweig (19), working with children from five to fourteen years of age, and Nowlis (17), using college men as *Ss*, found that their *Ss* tended to prefer of two tasks the one in which they had been successful rather than the one in which they had failed. Gebhard (9) demonstrated that experienced success, in general, was accompanied by a rise in attractiveness and experienced failure by a fall in the attractiveness of a task. In an earlier study Cartwright (4) found that anticipation of failure tended to reduce the attractiveness of an activity more than the actual experience of failure, and Gebhard (9), controlling expectation of success and failure, found that expected future success, in general, was accompanied by a rise in attractiveness and expected future failure by a fall in attractiveness. In another study Gebhard also demonstrated the importance of preperformance expectation of success and failure, concluding that "attractiveness rose when experienced success followed expected failure and fell when experienced failure followed expected success" and "attractiveness rose when experienced success followed expected success and also rose some-

[1] This paper is a dissertation submitted in partial fulfillment of the requirements of the degree of Doctor of Philosophy, in the Department of Psychology of the University of Pennsylvania. The writer is greatly indebted to Dr. Francis W. Irwin for his valuable suggestions and criticisms during the course of this investigation and is similarly grateful to Dr. Mildred Gebhard and Dr. Mortimer Garrison for their help. He sincerely appreciates the cooperation of the administration and teachers in the Fel Public School of Philadelphia from which *Ss* were obtained.

what when experienced failure followed expected failure" (8, p. 412).

The experiment reported in this paper was designed to investigate the effects of attainment and nonattainment, as well as the effects of other variables suggested by the studies dealing with activities, upon the attractiveness of goal objects. Before presenting the hypotheses, in order to prevent confusion, it is important to emphasize that goal objects are dealt with in this experiment. Lewin (14) points out that frequently the material object itself is not the final goal of the individual, but rather the goal may be an action or state, as the eating of food or the possessing of the object. The writer is using the terms goal and goal objects interchangeably, meaning by these terms objects which individuals strive to obtain. Three hypotheses concerned with three variables, attainment, expectation of attainment, and merit-involvement, were advanced to be tested by the experiment, and these with a discussion of them follow.

1. *Attainment and Nonattainment*

Hypothesis: Nonattainment of a goal object will tend to be accompanied by a decrease in the attractiveness of the goal object, whereas attainment of a goal object will tend to be accompanied by an increase in its attractiveness.

When an individual fails to obtain a goal object that he has been seeking, a situation referred to as frustration exists. The term frustration is commonly used to refer to instances where the satisfaction of a need is blocked, interfered with, or simply not available. Various responses in such situations have been studied and described in the psychological literature (2, 7, 16, 22). As is well known, these responses may involve general aggression, random movements, regression, or a specific reaction against the part of the environment considered responsible for the frustration. Another response that has been clinically noted is the reaction that personality theorists have called the "sour grapes mechanism." This behavior involves the re-evaluation of a goal which an individual has not been able to attain, so that the nonattained goal does not seem as important nor as valuable as it originally did. In the present experiment, by using a ranking procedure, *S*s were given the opportunity to re-evaluate the attractiveness of a goal object, and the above hypothesis predicts that under the condition of frustration there will be a decrease in the attractiveness of the goal object.

Much more attention has been devoted to frustration than to its counterpart, satisfaction, the situation when an individual has attained his goal. When a goal object is attained, one might expect reactions in this situation to be contrary to responses following nonattainment. Thus, passivity, the lack of random movements, absence of regression, and, according to the hypothesis of this experiment, a re-evaluation that will result in an increase in a goal object's attractiveness are the behavior concomitants of satisfaction. Furthermore, when a goal object is obtained, its status changes from nonbelonging to belonging, and it can be used to satisfy other individual needs. Dealing with stated preferences, Irwin and Gebhard (11) found that children tended to prefer objects that were to be given to them rather than objects that were to be given to others. With attainment of a goal object both satisfaction and possession occur, and it seems likely that a re-evaluation of the object's attractiveness will result in increased attractiveness.

2. *Expectation and Nonexpectation*

Hypothesis: Expectation of the attainment of a goal object will tend to be accompanied by an increase in attractiveness, whereas the expectation of nonattainment of a goal object will tend to be accompanied by a decrease in its attractiveness.

It seems reasonable to assume that an individual who expects to attain a goal object will perceive himself to be closer to the object than an individual who does not expect to attain the object. Although Wright (23) in one of the early studies in the area maintained that a barrier enhances the value of a goal object, it has been demonstrated in a number of experiments that the closer an individual is to an object either spatially or temporally the more attractive it is, and the repetition of Wright's work by Child and Adelsheim (6) produced results generally opposite to his results. In another investigation Child (5)

found that children at the age level of those used in this experiment generally preferred of two objects the one that was easier to obtain. In two studies using time as a variable Irwin and his students (10, 12) demonstrated that *Ss* tended to prefer an object that they were to receive immediately rather than one they would receive later. Experimenting with animals, Brown (3) found that rats restrained at a point near a goal exerted a greater amount of force against a restraining instrument than did rats restrained at a point more remote from the goal. The evidence of the above studies seems to support in several dimensions the principle that the attractiveness of a goal object increases the nearer the goal object is to the *S*. In accordance with this principle, one would expect a goal object to increase in attractiveness when an individual discovers it is easy to attain (or he becomes psychologically nearer to attaining it) and to decrease in attractiveness when he discovers it is difficult to attain (or he becomes psychologically farther away from attaining it).

3. *Merit-Involvement and Nonmerit-Involvement*

Hypothesis: Reactions to attainment and nonattainment, when an individual believes that attainment depends upon his merit, will be intensified.

The third hypothesis is based upon the assumption that an *S* will be more ego-involved when attainment of a goal object depends upon his ability (merit-involvement) than when attainment depends upon chance (nonmerit-involvement). Because of the presumably greater ego-involvement in the merit-involved situation, it seems reasonable to suppose that reactions to attainment and nonattainment will be more intense under merit conditions than under nonmerit conditions. The specific reactions to attainment and nonattainment were predicted by the first hypothesis. It is now further predicted that the increase in the attractiveness of a goal object following attainment will be greater when the individual believes that attainment depends upon his merit than when he believes it depends upon chance, and that the decrease in attractiveness will be greater following nonattainment under the merit-involved conditions than under the nonmerit-involved conditions. The summaries of the relevant experimental literature by Allport (1) and by Sherif and Cantril (21) point out significant behavioral differences between ego-involved and nonego-involved groups, and because this variable is frequently related to goal directed behavior, its effect through merit-involvement on the attractiveness of goal objects is worthy of study.

EXPERIMENTAL DESIGN AND PROCEDURE

With conditions of attainment and nonattainment, two conditions of expectation, and two of merit, there were possible eight combinations of experimental conditions and the experiment was planned according to the principles of factorial design. Twenty *Ss* were assigned randomly regardless of sex to each of the conditions, making a total of 160 *Ss*. Public school children in the second and third grades, ranging in age from six to nine years of age, were used. Of the 160 *Ss*, 84 were boys and 76 were girls.

Nine different goal objects were used in the experiment. These were toys selected from a dime store, each costing ten cents. They were a bolo-ball, rubber ball, bag of marbles, boat, gun, bag of balloons, crayons, horn, and harmonica. The objects were arbitrarily selected, their usefulness being demonstrated in preliminary experimentation, which indicated that no object was so attractive or unattractive that it was uniformly liked or disliked.

When an *S* entered the experimental room, he was shown the nine objects and was asked to rank them in the order which he liked them. The most preferred object was assigned a rank of one by the experimenter, the next most preferred object a rank of two, etc. The object given the middle rank (fifth in the series of nine) was used by the experimenter as the toy which the *S* could attain by playing a pegboard game, the directions for which depended upon the various combinations of experimental conditions. The particular board used was 7½ inches by 7½ inches and was subdivided into 36 squares. There were six squares of each of the following colors: red, green, blue, yellow, white, and black. In the center of each square there was a hole one-half inch in diameter into which pegs two and one-half inches long fitted. Two small gray boxes were also used. In one box there were approximately 50 small cards, on each of which there was a small colored square that corresponded to one of the various colors on the pegboard. In the other box were slips, on each of which was a symbol that *Ss* were led to believe determined which of several games they would play. Although the directions varied, the actual task was physically the same for all the experimental conditions.

According to the design, half of the *Ss* attained the goal object and half did not. The experimenter accomplished this by stopping one-half of the merit-involved group before they completed the task and

by knowing the location of the various colors in the box so that he could select a color that would enable an *S* to attain or not attain a goal object under the nonmerit-involved conditions as explained in the directions below. To vary expectation half of the *Ss* were informed that the game they were about to play was an easy one in which almost everybody got the prize, and the other half were informed that the game was a difficult one. To vary merit, half of the *Ss* were given instructions which indicated that attainment depended on their ability, and half were given instructions that indicated attainment depended upon chance factors.

After playing the game the *Ss* were asked to make a second ranking of the nine objects and to give monetary value to the critical object as well as to the object they ranked second in the postexperimental ranking. Preliminary work indicated that young children had vastly different scales of values. One child might state that an object would cost a dollar while another child would state that the cost of the same object was five cents. By using the two objects it is possible to get a relative value of the critical object. The second object was selected because preliminary work indicated that it was generally above the second ranking of the critical object and thus generally two value judgments could be obtained. Changes in the attractiveness of the goal object were determined by comparing the second rankings of the goal objects with the first ranking which, as mentioned above, was always five. In addition, by using each child's estimate of the cost of the second object in the second ranking as a standard, the relation of the estimated value of the critical object to this standard could be compared for the various groups.

A personality questionnaire [2] was administered at the completion of the experiment to entire classes from which *Ss* had been used with the hope that deviant responses in the experiment might be related to personality adjustment. The shortcomings of the questionnaire technique were recognized, but it was hoped that there would be enough high and low adjustment scores to investigate the reactions of "well-adjusted" and "poorly-adjusted" individuals. Zander (24) has demonstrated that emotionally unstable persons react aggressively to a smaller degree of frustration and interference than do stable persons. Using a personality questionnaire offered the opportunity to compare further the reactions of subjects with different adjustment ratings.

After an informal greeting of the child, the experimenter gave the following instructions from memory.

Establishment of First Ranking:

> I want to find out how you like certain toys. You see, I have nine things here. I want you to tell me which one you like most, which one you like next best, which one you like after that, and so on. Which one do you like best? Which one do you like next best? Which one do you like after that? (Repeat until all have been selected.)

Introduction to Game:

> I want you to play a game with me. It is possible for you to get a prize in this game. This is the prize you can get in the game we are going to play. I have different games; some are easy and some are hard. I'm going to draw a slip from the box to see which one you will play.[3] I will tell you how to play the game and how you can get the prize.

Establishment of Expectations: [4]

> 1. The game you will play is a pegboard game. It is very easy to get the prize in this game. Almost everybody who plays this one gets the prize.
> 2. It is very hard to get the prize in this game. Almost nobody who plays this one gets the prize.

Establishment of Merit Involvement:

1. Nonmerit.

> When I tell you to start you are supposed to place the small sticks in the holes [illustrating]. I am going to give you a certain amount of time and then tell you to stop. When I tell you to stop, you will be in one of these colored squares—red, green, or one of the others. I am then going to draw a slip from this box which has a lot of slips with different colors.
>
> a. If the color on the slip is different from the block in which you have put the last stick, you will get the prize. If it is the same color, you won't get it. Nothing you can do will help you. It all depends on what color I draw from the box and not on you. You see, there are a lot more colors than the one in which the last stick will be, so it looks as if you have a good chance of getting the prize. This is an easy game.
>
> b. If the color on the slip is different from the block in which you have put the last stick, you will not get the prize. If it is the same color, you will get it. Nothing you can do will help you. It all depends on what color I draw and not on you. You see, there are a lot more colors than the one in which the last stick will be, so it looks as if you don't have much chance of getting the prize. This is a hard game.
>
> Do you understand how to play?

2. Merit

> Actually, it depends on how fast you are. This is how you play the game. You are supposed to place the small sticks as fast as you can into the holes. (Illustrating.) I am going to give you a

[2] California Test of Personality, Primary Series, Grades, Kgn—3; Thorpe, Clarke, and Tiegs.

[3] This was to create the impression that there was a chance selection of games. Also at this point the experimenter asked *Ss* their names and ages as he arranged the game that had been drawn for the child.

[4] In reading the instructions it must be remembered that in order to establish expectations, half of the *Ss* were informed that it was a difficult game. Besides having different general statements about the game it was also necessary to have different statements under each of the conditions of merit involvement which were consistent with the conditions of expectations. These have been noted as (*a*) and (*b*) under the instructions for merit involvement.

certain amount of time and then tell you to stop. If you get all the sticks in before the time is up, you get the prize. If you don't get them in within the time allowed, you don't get the prize. It depends on how good you are whether or not you get the prize.

(To reinforce expectation one of the following was used.)

a. Most people usually are able to finish in the time allowed, so it will be easy for you to get the prize.

b. Most people usually are not able to finish in the time allowed, so it will be hard for you to get the prize.

When I tell you to start, you are to begin, and stop when I tell you to stop. Do you understand how to play?

Establishment of Attainment:

(To reinforce attainment or nonattainment at the end of the game one of the following was used.)

a. You got the prize.
b. You did not get the prize.

Establishment of Second Ranking:

I want to find out again how you like the prizes. People often change their minds about things. If you have changed your mind about how you like the prizes, which I showed you before, that is OK. In order to find out how you like all the prizes, I'm going to put back the one you got (or didn't get) with the others for a minute. Which one of these prizes do you like best this time? Which one do you like next best? Which one do you like after that?

Estimate of Cost:

[To find out what monetary value the children placed on the goal object, the following questions were asked after the second ranking had been established.]

How much do you think the goal object would cost if you wanted to buy it in a store? How much do you think the object ranked second would cost if you wanted to buy it?

[As they left the room, *S*s were asked not to talk about the experiment.[5]]

RESULTS AND DISCUSSION

Analysis of Measures of Attractiveness

It will be remembered that the goal object used for each *S* was the toy which had been ranked fifth in the *first* rankings of attractiveness. A value less than five in the *second* ranking indicates an increase in attractiveness and a value greater than five indicates a decrease in attractiveness. Table 1 shows the mean second rank of the goal object under each of the eight conditions, and lists the means for the conditions of attainment, nonattainment, expectation, nonexpectation, merit-involvement, and nonmerit-involvement. The mean rank of all of the second rankings was 5.0, a value that happens to be the same as the common rank of 5 for the goal objects in the first ranking.

TABLE 1

MEAN RANK OF ATTRACTIVENESS OF GOAL OBJECT ON NINE-POINT SCALE AFTER PERFORMANCE UNDER EXPERIMENTAL CONDITIONS

	Attainment		Non-Attainment	
	Expect Attainment	Expect Non-Attainment	Expect Attainment	Expect Non-Attainment
Merit	3.9	4.3	6.1	6.3
Non-Merit	4.3	3.7	5.5	6.1

Combined Means

Attainment	4.1	Expectation	4.9
Non-Attainment	6.0	Non-Expectation	5.1
Merit-Involvement	5.2		
Non-Merit-Involvement	4.9		

An analysis of variance was made to determine the statistical reliability of the differences resulting from the effects of the three experimental variables, attainment, expectation of attainment, and merit-involvement. It was clear that the attainment variable had a strong effect in the second rankings of the goal object, p being far below .01.[6] However, the conditions of expectation and merit-involvement and all the interactions showed no significant effects.

An examination of Table 1 shows that under each condition of attainment there was an increase in the attractiveness of the goal object and under each condition of nonattainment there was a decrease in attractiveness. Both the mean rank after attainment (4.1) and the mean rank (6.0) after nonattainment differ significantly from the initial ranking, as well as differing significantly

[5] Those that attained a goal object in the experiment were informed that the object would be brought to them at a later time if they would not tell anyone about the game. Those who did not attain an object were informed that they would be given another chance later to get a toy if they did not talk about the experiment. Answers to questions asked by the experimenter to some *S*s seemed to indicate that there was little factual communication about the details of the experiment. Some *S*s knew that they were going to play a game and could get a toy. After all *S*s had participated in the experiment, the goal objects were given to those who attained them and token toys were given to the others.

[6] An F-value of 35.3 was obtained. For the conditions of the experiment an F of 6.81 has a p-value of .01.

from each other ($p<.01$ in all three instances). Under conditions of expectation of attainment and expectation of nonattainment, the mean second rankings of the goal objects were 4.9 and 5.1 respectively. Neither of these values differs significantly from the initial ranking, nor do they differ significantly from each other. The difference between the mean second rank for conditions in merit-involvement (5.2) and the mean second rank for conditions not involving merit (4.9) is likewise not statistically reliable and neither of these means is significantly changed from the first ranking.

It is also possible to examine the relative frequencies of the instances where the second rank of the goal object is higher, lower, or unchanged from the initial rank of five. If a 3 x 2 table is arranged, subclassified as to the frequency of rankings showing increased, decreased, or unchanged attractiveness and as to attainment or nonattainment of the goal object, a chi-square test yields a p-value of less than .01. The difference is in the same direction as that obtained from the mean rankings; when the goal object was obtained more rankings show increased attractiveness than either unchanged or decreased attractiveness and when the goal object was not obtained, the rankings show more frequently a decrease in attractiveness than no change or increased attractiveness. A 3 x 2 chi-square test between changes in attractiveness and expectation, and a similar test between changes in attractiveness and merit-involvement yields in each case a p-value $>.10$.

An analysis of the monetary ratios (estimated cost of object ranked second in post-game ranking divided by estimated cost of goal object) [7] revealed, as might be expected, that the distributions were skewed. However, a logarithmic transformation of the ratios brought the distribution close to symmetry. The mean logarithm of the ratios for all the conditions of attainment is .918 and the mean logarithm of the ratios for the nonattainment conditions is .991, the difference being statistically significant ($p<.05$).[8] The differences between the means for the other conditions, expectation and merit-involvement, are not statistically reliable. The results thus indicate that a greater relative monetary value was placed upon the goal object following attainment than following nonattainment.

The results of the second ranking and of the monetary values of the goal objects support the hypothesis that nonattainment of a goal object will tend to be accompanied by a decrease in the attractiveness of the goal object, while attainment of a goal object will tend to be accompanied by an increase in its attractiveness. These results confirm Lewin's statement quoted in the introduction of this paper and demonstrate the direction of changes in valences of goal objects after attaining and not attaining them. In an attempt to understand their behavior further, the experimenter asked children who had large differences between the first and second ranks of the goal object, "Why do you like the (name of goal object) better (or not as much) this time?" Certain replies seem to support the theoretical structure presented in the first part of this paper in defense of the hypothesis relating to attainment and nonattainment. While most of the children questioned stated that they did not know why they had changed their minds about the object, some significant answers were: "Because I didn't get it," "It made me lose," "Because I won it," and "It gave me bad luck."

The hypothesis that expectation of attainment of a goal object would lead to an increase in its attractiveness whereas the expectation of nonattainment of a goal object would lead to its decrease in attractiveness, and the hypothesis that there would be an interaction between merit-involvement and attainment, were not supported by the results of the experiment. If it is assumed that the various conditions of expectation and merit-involvement existed for *S*s in the experiment, it is necessary to conclude that expectation and merit-involvement have no influence upon the attractiveness of goal objects for young children six to nine years of age. It is possible, however, that the procedures used did not generally produce the desired conditions, and before rejecting the hypotheses

[7] If the goal object itself was ranked second, only one estimate could be obtained, and these instances were not used in computing the monetary ratios.

[8] Because the direction of this difference was predicted by the first hypothesis of the experiment, the experimenter used a one-tail test of its significance.

an examination of the experimental procedures is advisable. Admittedly, this examination is *ad hoc,* but its use must be weighed against the necessity of otherwise rejecting not only the specific hypotheses of the experiment but some of the assumptions related to these hypotheses. For example, when the expectation hypothesis of this investigation is rejected, it will also be necessary to conclude either that expectations are not related to goal distances or that goal distances are not related to goal attractiveness, a conclusion that would be inconsistent with the data of several experiments mentioned in the discussion of the expectation hypothesis. Likewise, if the merit-involvement hypothesis is rejected, it would also be necessary to conclude either that different degrees of ego-involvement do not accompany different degrees of merit-involvement or that, since it has been demonstrated that goal attractiveness is related to frustration and satisfaction, different degrees of ego-involvement are not accompanied by different reactions to frustration and satisfaction. Because of the consequences of a complete rejection of the hypotheses of the experiment, it seems prudent first to examine the adequacy of the experimental conditions.

With respect to the conditions of expectation, verbal information about the game may not have been sufficient to establish either expectations or different degrees of expectation for the children. Child (5), for example, found that adding the cue words "hard" and "easy" to attainment instructions for his subjects made no significant difference in their choice of goal objects. In Gebhard's experiment (9), on the other hand, instructions appear to have been adequate to establish genuine differences in expectation. Gebhard, however, worked with college students who had had many more verbal experiences than *Ss* in the present experiment. In a study of children's stated expectations Marks (15) demonstrated that when there was only a .5 probability of obtaining a desired object children stated, nine out of ten times, that they expected to obtain the object. Because the goal object in the present experiment was a desirable toy, it is possible that most *Ss,* even when they were informed that attainment was extremely unlikely, expected to obtain the toy. While this possibility may account for the failure to obtain differences between the two experimental conditions of expectation, it does not, without assuming some general counteracting effect, account for the fact that there was no over-all rise in the attractiveness of the goal object in the second rankings. If most of the subjects expected to attain the goal object, then according to the original hypothesis there should have been a general rise in the attractiveness.

With respect to the conditions of merit-involvement the *Ss* in the experiment seemed to understand the chance and merit factors used in investigating the third hypothesis, and it does not appear likely that the instructions were too complicated. Several *Ss* voluntarily commented "I know" when the experimenter informed them that attainment depended upon how good they were or that attainment depended upon chance. However, the attempt to get the toy may have been the crucial ego-involving factor in the experimental situation, and nonmerit-involved children, even though they were in a chance situation, may have become sufficiently ego-involved in the game that there was no difference in the psychological involvement of the *Ss* under the two merit conditions. Problems of "ego" development also seem relevant to the task of establishing degrees of ego-involvements. Early stages of "ego" development include awareness and recognition of the self and interpretation of the self-environment relationships. At about the age of six a child begins to set up standards and to form value judgments about his behavior (21). It is possible that *Ss* in the present experiment had not reached the stage of "ego" development where conditions of individual merit in this experiment would lead to different degrees of ego-involvement. Rosenzweig (19), in his study on the repetition of successful and unsuccessful activities, reported that children below the median in a group with ages ranging from five years and six months to fourteen years and eight months had lower ratings by teachers on "pride" than children above the median. Sanford (20) suggests that the discrepancies between his and Rosenzweig's results (18) may be partly due to the fact that there were different degrees of ego-involvement between the two groups of *Ss* in their experiment. *Ss*

in Sanford's experiment were older than Rosenzweig's, ranging from seven to fifteen years of age, and all had mental ages well above eight years and five months. Over two-thirds of the children in the present experiment were under eight years of age and all were under ten. Thus, one explanation for the failure to obtain significant differences under merit conditions may be the general level of "ego" development of children within the age range of the experimental group, i.e., *Ss* may not have reached the necessary level of "ego" development for a differentiation of ego-involvements, related to the merit-involved conditions, to occur. The narrow age range of the experimental children does not permit an adequate analysis of results by age groups to test this explanation.

Another noteworthy consideration is the indication that the attainment of the goal object was of prime importance for the children in this experiment. The analysis of variance reported above revealed the extreme significance of the attainment variable. It is possible that the influence of this variable was so psychologically important that its effect tended to depress other experimental effects. The second ranking was obtained immediately following attainment or nonattainment when the reactions to frustration and satisfaction were likely to occur.

Sex Differences

Analysis of the data with respect to sex of the subjects showed some interesting significant differences. There were 42 boys and 38 girls under each of the two conditions of attainment. After attainment the boys' mean rank of the goal object was 4.5 and the girls' mean rank was 3.6, while following nonattainment the boys' mean rank was 6.3 and the girls' was 5.4. Thus, the goal object, following both attainment and nonattainment, was more attractive for girls than for boys, the difference between the mean values for boys and girls in each case being statistically significant ($p<.05$ in each case). While the data of the present experiment cannot explain these differences, it is of some importance to find them at the age level of these *Ss*. It is possible that, for some reason, goal objects in a second ranking generally may be more attractive to girls than to boys. It is also possible that reactions to satisfactions are more intense for girls than they are for boys and that reactions to frustrations are more intense for boys.

Results of the Personality Inventory

To ascertain the effects of personality adjustment upon the responses of the *Ss*, "well-adjusted" and "poorly-adjusted" groups were designated. Of 154 *Ss* who participated in the experiment and also marked the inventory there were, using the published norms of the test, 40 cases at the 75th percentile or above (good adjustment) and only 19 cases at the 25th percentile or below (poor adjustment). These values were arbitrarily selected to designate well-adjusted and poorly-adjusted groups. Higher and lower adjustment scores would have reduced the number of subjects so that comparisons would have been impossible. It is apparent that the distribution of percentile scores contained more high adjustment scores than low adjustment scores.

For the well-adjusted group the mean rank following attainment is 3.9 and following nonattainment 6.1, and for the poorly-adjusted group these mean ranks are 4.9 and 5.4 respectively. The increase in the attractiveness of the goal object following attainment and the decrease in attractiveness following nonattainment are statistically reliable for the well-adjusted group ($p<.01$ in each case), whereas the changes in the attractiveness for the poorly-adjusted group are relatively small and are not statistically reliable ($p>.10$ in each case).

When the differences between the two adjustment groups were analyzed, it was found that the difference between the mean second ranks of the two groups following attainment only approaches statistical reliability (p between .05 and .10), and the difference between the mean second ranks following nonattainment is not statistically reliable ($p>.10$). By noting the number of cases that showed an increase in the attractiveness of the goal object following attainment and the number of cases that showed a decrease following nonattainment, it is also possible to determine for the two groups the percentage of reactions that were in the directions predicted by the attainment hypothesis. This

treatment of the results revealed that following attainment 60 per cent of the well-adjusted group behaved in the manner predicted by the attainment hypothesis, whereas only 22 per cent of the poorly-adjusted group behaved in the predicted manner, a p-value of $<.05$ indicating that the difference between the two percentages is statistically reliable. Following nonattainment 67 per cent of the well-adjusted group and 50 per cent of the poorly-adjusted group behaved as the hypothesis predicted, the p-value of $>.10$ indicating that the difference is not significant. Considering both conditions of attainment, 63 per cent of the well-adjusted group and 37 per cent of the poorly-adjusted group behaved in the manner predicted by the attainment hypothesis, the difference only approaching statistical reliability (p between .05 and .10). A similar analysis of the mean ranks and the percentage of cases in the direction of the hypotheses revealed no noteworthy differences between the two adjustment groups under the conditions of expectation and merit-involvement. Because of the generally low statistical reliability of the differences between the results of the two adjustment groups under the attainment conditions, no definite conclusions should be drawn until these results are further confirmed.

SUMMARY AND CONCLUSIONS

The experiment reported in this paper was designed to investigate the effects of three variables, attainment, expectation of attainment, and merit-involvement, upon the attractiveness of goal objects. To study these effects 160 public school children, 84 boys and 76 girls with an age range of six to nine years, performed a task to try to attain a goal object. With conditions of attainment and nonattainment, two conditions of expectation of attainment, and two of merit-involvement in the experiment there were eight combinations of experimental conditions and 20 *S*s were assigned randomly regardless of sex to each of the conditions. By using a ranking procedure the relative attractiveness of toys was determined, the particular goal object which an *S* tried to attain always being the object ranked fifth in an initial ranking of nine objects. Following performance of the tasks under the various experimental conditions a second ranking of the nine objects, as well as a relative monetary value of the goal object, was obtained. Changes in the attractiveness of the goal object, as indicated by changes in the two rankings and the relative monetary value, were analyzed to determine the effects of attainment, expectation, and merit-involvement. By using a personality questionnaire well-adjusted and poorly-adjusted groups were designated and the responses of these two groups were also compared.

The results of the study are summarized in the following statements.

1. Attainment of a goal object by children was generally accompanied by an increase in the attractiveness of the object and nonattainment was generally accompanied by a decrease in attractiveness.

2. Instructions to the effect that a goal object was likely to be attained or was not likely to be attained had no effect upon the attractiveness of the goal object following attainment and nonattainment.

3. Instructions used to create experimental conditions of merit-involvement and non-merit-involvement had no effect upon the attractiveness of a goal object following attainment and nonattainment.

4. Following both attainment and nonattainment the critical goal objects were significantly more attractive to girls than to boys.

5. Following attainment and nonattainment, the differences between the responses of well-adjusted and poorly-adjusted groups were inconclusive. However, there was a significant rise in the attractiveness of a goal object following attainment and a significant decrease in attractiveness following nonattainment for well-adjusted *S*s, while there were no significant changes in the attractiveness of a goal object following either attainment or nonattainment for poorly-adjusted *S*s.

REFERENCES

1. ALLPORT, G. W. The ego in contemporary psychology. *Psychol. Rev.*, 1943, **50**, 451–478.
2. BARKER, R., DEMBO, T., & LEWIN, K. Studies in topological and vector psychology: II. Frustration and regression. *Univ. Iowa Stud. Child Welf.*, 1941, **18**, No. 1.
3. BROWN, J. S. Gradients of approach and avoidance responses and their relation to the level of motiva-

tion. *J. comp. physiol. Psychol.*, 1948, **41**, 450–465.

4. CARTWRIGHT, D. The effect of interruption, completion, and failure upon the attractiveness of activities. *J. exp. Psychol.*, 1942, **31**, 1–16.
5. CHILD, I. L. Children's preference for goals easy or difficult to obtain. *Psychol. Monogr.*, 1946, **60**, No. 280.
6. CHILD, I. L., & ADELSHEIM, E. The motivational value of barriers for young children. *J. genet. Psychol.*, 1944, **65**, 97–111.
7. DOLLARD, J., DOOB, L. W., MILLER, N. W., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression*. New Haven: Yale Univ. Press, 1939.
8. GEBHARD, MILDRED E. Changes in the attractiveness of activities: The effect of expectation preceding performance. *J. exp. Psychol.*, 1949, **39**, 404–413.
9. GEBHARD, MILDRED E. The effect of success and failure upon the attractiveness of activities as a function of experience, expectation, and need. *J. exp. Psychol.*, 1948, **38**, 371–378.
10. IRWIN, F. W., ARMITT, F. M., & SIMON, C. W. Studies in object-preferences. I: The effect of temporal proximity. *J. exp. Psychol.*, 1943, **33**, 64–72.
11. IRWIN, F. W., & GEBHARD, MILDRED E. Studies in object-preferences: The effect of ownership and other social influences. *Amer. J. Psychol.*, 1946, **59**, 633–651.
12. IRWIN, F. W., ORCHINIK, C. W., & WEISS, J. Studies in object-preferences: The effect of temporal proximity upon adults' preferences. *Amer. J. Psychol.*, 1946, **59**, 458–462.
13. LEWIN, K. Behavior and development as a function of the total situation. In L. CARMICHAEL (Ed.), *Manual of child psychology*. New York: Wiley, 1946, 791–844.
14. LEWIN, K. *Principles of topological psychology*. New York: McGraw-Hill, 1936.
15. MARKS, ROSE W. The effect of probability, desirability, and "privilege" on the stated expectations of children. *J. Pers.*, 1951, **19**, 332–351.
16. MURPHY, G. *Personality, a biosocial approach to origins and structure*. New York: Harper, 1947.
17. NOWLIS, HELEN H. The influence of success and failure on the resumption of an interrupted task. *J. exp. Psychol.*, 1941, **28**, 304–325.
18. ROSENZWEIG, S., & MASON, G. An experimental study of memory in relation to the theory of repression. *Brit. J. Psychol.*, 1934, **24**, 247–265.
19. ROSENZWEIG, S. Preference in the repetition of successful and unsuccessful activities as a function of age and personality. *J. genet. Psychol.*, 1933, **42**, 423–441.
20. SANFORD, R. N. Age as a factor in the recall of interrupted tasks. *Psychol. Rev.*, 1946, **53**, 234–240.
21. SHERIF, M., & CANTRIL, H. *The psychology of ego-involvements*. New York: Wiley, 1947.
22. SYMONDS, P. M. *Dynamic psychology*, New York: Appleton-Century-Crofts, 1949. Pp. 42–94.
23. WRIGHT, H. F. The influence of barriers on strength of motivation. *Contr. Psychol. Theory*, I, 1937, 1–143.
24. ZANDER, A. F. A study of experimental frustration. *Psychol. Monogr.*, 1944, **56**, Whole No. 256.

Received March 23, 1951.

THE PERCEPTUAL CONSEQUENCES OF SUCCESS AND FAILURE[1]

LEO POSTMAN AND DONALD R. BROWN
University of California

THE recent emphasis on motivational and personality factors in perception has called attention to the selectivity of perceptual organization. There is a common thread that runs through the studies concerned with the systematic effects on perception of physiological needs (7, 9, 21), personal values (17, 20, 23), social attitudes and personality structure (5, 6, 10, 11). Under a variety of experimental conditions and with widely diverse stimulus materials perceptual response has been shown to be selective, responsive to the demands of motivation, and modifiable through experience.

The selectivity of perception is not necessarily instrumental to wish-fulfillment or tension reduction. In connection with the investigation of motivational and personality variables there is a temptation to think of perception as "autistic" (13)—serving the needs and wishes of the organism within the limits allowed by the stimulus conditions. Such may indeed be the case in some situations, but the assertion that perceptual selectivity reflects the motives and past experiences of the individual has higher generality: the possibility is left open that perceptual selectivity may favor negatively valued and threatening as well as positively valued and need reducing objects (3). The characteristics of objects with respect to which selectivity operates may, indeed, fall along dimensions altogether different from wish-fulfillment and wish-denial. At this stage of development, the variables relating perception to personality need not necessarily come from a particular theory of personality, such as psychoanalysis. The latter would, of course, focus attention on the wish-fulfilling and repressive functions of perception. It may be possible to state the determinants of selectivity in more general terms, so that existing personality variables would appear as special cases. The analysis of perceptual selectivity in terms of concepts derived from learning theory may provide such an approach.

Borrowing a term from the theory of discrimination learning, we have suggested elsewhere that the processes of perceptual organization and reorganization may be conceptualized in terms of the arousal and confirmation of hypotheses (2, 14). Hypothesis, as an intervening variable, denotes a predisposition of the perceiver to transform and organize stimulus cues from the environment in specific ways. These stimulus cues include not only characteristics of specific objects but also cues defining the context or situation in which the objects appear. An hypothesis is conceived as determined by both the properties of the sensory system ("autochthonous" factors) and other characteristics of the perceiver, such as his past experiences and motivational state ("directive" factors). Operationally, then, hypothesis is anchored in manipulations of the sensory input *and* directive factors on the one hand and in the perceiver's systematic responses (discriminations, matches, verbal reports) on the other.

Stimulus cues serve to arouse, confirm, or deny hypotheses derived from past experience. In the case of complex stimuli and under difficult conditions of discrimination, the achievement of a stable perceptual organization may involve a cyclical process of trial and check of hypotheses (24). An hypothesis aroused by an incomplete set of cues may or may not be confirmed by subsequent information from the environment. If the hypothesis is denied, it may shift and shift again until confirmation brings a stable perceptual organization. The more readily an hypothesis is aroused and confirmed, the more speedy and effective is perceptual organization. The selectivity of perception reflects the greater readiness with which some hypotheses rather than others are aroused and confirmed.

The readiness with which an hypothesis is aroused and confirmed depends on the

[1] We gratefully acknowledge the cooperation of Dr. Donald W. MacKinnon in making available to us the facilities of the Institute of Personality Assessment and Research of the University of California.

strength of the association between stimulus cues (including situational cues) and hypotheses. The higher the degree of this association, the *stronger* the hypothesis. Hypothesis strength is measured by the amount of stimulus cues necessary for the arousal and confirmation of the hypothesis: the stronger the hypothesis, the less the amount of stimulus cues required to arouse and confirm it.

The laws governing the acquisition of hypothesis strength are the laws of associative learning. We have suggested elsewhere (14) that recency and frequency of past confirmation and emphasis resulting from motivational relevance are among the determinants of hypothesis strength.

It follows from this analysis that perceptual selectivity need not necessarily favor wish-fulfilling objects over wish-denying ones. The constellation of cues in any particular situation may be more strongly associated with one or the other type of hypothesis. Whichever hypotheses are stronger will determine the observed selectivity. If our analysis is correct, it should be possible to produce perceptual selectivity for either wish-fulfilling or wish-denying objects by experimental manipulation of situational cues. Specifically, if situational cues can establish a context of success in Situation A and a context of failure in Situation B, there should be perceptual selectivity for success-related (wish-fulfilling) objects and symbols in Situation A and for failure-related (wish-denying) objects and symbols in Situation B.

The following experiment was designed to show that systematic variations in situational context do lead to predictable shifts in sensitivity to wish-fulfilling and wish-denying symbols. With words connoting success and failure as stimuli, the following prediction was tested: an experimentally induced context of failure results in selective sensitivity to failure words; an experimentally induced context of success results in selective sensitivity to success words.

THE EXPERIMENT

The experiment consisted of two procedures: (*a*) establishment of a situational context of success or failure, and (*b*) measurement of perceptual sensitivity to symbols of success and failure (words).

Part I: Situational Context of Success and Failure

To establish a context of success or failure, the *level of aspiration* situation was used. The procedure was so arranged that some *S*s attained or exceeded most of their stated goals in an experimental task while other *S*s did so only rarely. Thus, contexts of success or failure could be created under standard conditions.

Experimental task. The task in which *S*s succeeded or failed and for which their level of aspiration was measured was a span of apprehension test. A series of 15 slides was shown on each of which a set of 12 symbols (capital letters, small letters, and numbers) appeared. The spatial arrangement of the symbols varied considerably from slide to slide. Each of the slides was exposed tachistoscopically for 1 sec., and *S*'s task was to report as many of the symbols as he could. The number of symbols which can be recognized under these conditions rarely exceeds six.

Instructions. There were three groups of *S*s: a Control Group, a Success Group, and a Failure Group. Members of the Control Group received the following instructions at the beginning of the experiment:

> I am interested in measuring the visual span of apprehension for a representative sample of college students. In order to do this, I shall expose on the screen in front of you a group of letters and numbers and I want you to tell me after each exposure which of the symbols you were able to see. Now I shall show you a few sample slides to give you an idea of the task. (*E* exposes three slides.)
>
> Is everything all right? OK, just relax and look at the screen whenever I say "Ready," and shortly after, the slide will be exposed. Remember, do as well as you can.

The following additional instructions were given to the Success Group and the Failure Group:

> Since you now know how well you were able to do on the practice slides, let us make the task a little more interesting by having you state before each exposure how well you feel you will be able to do. Of course, how well you do on this kind of task depends on a number of things. The most important of these are how well you can concentrate, how efficiently you can respond to changing symbols in the environment, and your speed of reaction. Eyesight, with fairly normal correction, doesn't seem to make a difference. In other words, the things that make for success in this task are the same things that are involved in good driving, accurate shooting, and other perceptual-motor tasks. Make your estimate by using percentage ranks which are based on a curve similar to a course-grade curve which you are no doubt acquainted with. In other words,

before each exposure you tell me what percentage of our previous subjects will fall below your score. For example, if you say 70, that means that 70 per cent of our subjects were below you, while 30 per cent were above you. Or if you say 45, that means that 45 per cent were below and 55 per cent were above your score. After you estimate your percentage score, I shall expose the slide and you tell me as in the practice trials what you were able to see. Then I shall check your actual score for that slide in our tables which have been standardized on a large population of students and tell you how close you came in your estimate. For your convenience I am giving you a piece of paper to record your scores, estimates, and the values I give you from the table under the proper headings. This will enable you to keep track of how well you are doing. I might add that these slides vary in difficulty, and I shall begin by exposing some of the easier slides and go on to more difficult ones.

Measurement of level of aspiration. Following the instructions, the series of 15 slides was exposed. Members of the two experimental groups announced their estimated scores *as percentile ranks* prior to exposure of each slide and were given their score after each exposure, again as percentile ranks. Before announcing the score, *E* went through the motions of looking up the *S*'s standing in a table of norms. Use of the percentile ranks had, of course, two advantages: first, a standard prearranged sequence of success and failure trials could be given independently of *S*'s performance; second, reference to norms established on a population of *S*'s peers probably emphasized the experience of success and failure.

The ranks announced to members of the Success Group were such that *Ss* exceeded their estimates on 11 out of the 15 trials. On four trials (nos. 3, 6, 8, and 11) the achievement announced to *Ss* barely reached, or fell short of, their estimates. Similarly, the percentile ranks announced to the members of the Failure Group were such that their achievement fell short of their estimates on 11 out of 15 trials. On four trials (again nos. 3, 6, 8, and 11) the failure *Ss* did reach or exceed their estimates. In each case, the four atypical trials were interspersed in order to give a more realistic appearance to the total situation.

Members of the Control Group were tested with the same 15 slides. The control *Ss* did not, of course, make any estimates of their performance, nor were they given any knowledge of results. They had, however, the same amount of practice in tachistoscopic recognition as did the experimental *Ss*.

Part II: Test of Perceptual Sensitivity to Success and Failure Symbols

Stimuli. Immediately following the span of apprehension procedure a series of 24 words was presented to all *Ss* for tachistoscopic recognition. There were 12 critical, motivationally relevant words and 12 control words. Of the 12 critical words, four were "goal words" connoting success, four were "deprivation words" connoting failure, and four were "instrumental words" connoting instrumental striving.[2] For each of the critical words there was a control word. Critical words and control words were matched in familiarity by means of the Thorndike-Lorge Semantic Count (22). The frequency of usage of critical words and corresponding control words is almost identical in every case. Not only were critical words and control words matched for familiarity but the three subgroups of words were similar in frequency of usage as well. Table 1 lists the pairs of critical and control words and their frequency of usage. The words were presented in random order so that different types of words followed each other in a haphazard order.

TABLE 1

LIST OF STIMULUS WORDS

CRITICAL WORDS	FREQUENCY OF USAGE *	CONTROL WORDS	FREQUENCY OF USAGE *
	Deprivation Words		
failure	262	similar	264
unable	239	ancient	228
obstacle	56	cupboard	55
defect	47	muffle	47
	Goal Words		
excellent	345	railroad	367
succeed	264	language	255
perfection	91	ingredient	94
winner	74	horizon	76
	Instrumental Words		
improve	269	presence	277
achieve	249	delicate	248
strive	56	tropic	57
compete	51	barber	50

* L-Count in *The Teacher's Word Book* (22), giving frequency of usage in approximately 4½ million words.

Determination of recognition thresholds. The stimulus words, typed in capital letters, were presented on 2 x 2 lantern slides. The slides were flashed on a ground-glass screen from a projector equipped with photographic shutter. Two parallel lines, exposed by means of a second projector, provided the fixation area within which the stimuli appeared. Speed of exposure was kept constant at 1/100 sec. The threshold of recognition was determined by increasing the brightness of the exposure.

At the beginning of the procedure, the approximate threshold of the *S* was determined by means of a practice slide. Exposures of all subsequent slides were then begun about 15 steps below the estimated threshold and continued until correct

[2] This classification of need-related words was suggested by McClelland and Liberman (10). Many of our stimulus words were taken from their list.

recognition of the word. Each step represented a change of 1.3 volt. The starting level of illumination thus varied somewhat from *S* to *S* but remained constant for a given *S* throughout the procedure. A measure of the recognition threshold was provided by the number of exposures required for correct recognition.

Instructions. Before exposure of the first slide, *Ss* in all groups were given the following instructions:

> I am now going to expose words instead of unconnected letters and symbols (as in the first part). I shall expose the same word over and over until you get it correctly, and then we shall go on to the next word. When I say, "Ready," look at the parallel lines on the screen and shortly thereafter the word will be shown between the lines. Tell me after each exposure what the word is, or what you think it is. Do not be afraid to guess if you are not absolutely sure that you saw it. Just relax and try to do as well as you can.

Subjects. The *Ss* were male undergraduate and graduate students at the University of California who had volunteered to take part in the experiment. The members of the *Failure Group* were tested under rather special conditions. These *Ss* were undergoing a series of intensive tests and interviews at the Institute of Personality Assessment and Research of the University. Our experiment was one of the procedures which the *Ss* underwent during the assessment. It was hoped that the context of intensive personality assessment would increase these *Ss*' sensitivity to failure. The *Ss* in the other two groups were volunteers from various classes in the University. Each *S* was tested individually. The experimental procedure took about one hour. The numbers of cases in the groups are uneven. There were 39 *Ss* in the Failure Group (total assessment group) and 14 *Ss* each in the Success Group and Control Group.

RESULTS OF THE EXPERIMENT

Level of aspiration behavior. The level of aspiration behavior of the experimental groups suggests that contexts of success and failure were, indeed, created for these *Ss* during the first part of the procedure. In Fig. 1 *Ss*' estimates of their performance and the achievement scores announced to them are plotted for the 15 trials of the experimental task (span of apprehension test). Both estimates and achievement scores are in terms of percentile ranks, evaluating *S*'s performance against fictitious population norms. Initial estimates by members of both groups are close to a percentile rank of 50. Thereafter the two groups steadily diverge. The estimates of the Success Group kept increasing until they finally reached an average percentile rank of 84.8. Since members of this group were allowed to reach or exceed their goals on most of the trials, the achievement scores announced to them rose concurrently. The estimates of the Failure Group gradually dropped down to an average percentile rank of 30.8 on the last trial. The members of this group failed to reach their estimates on most of the trials, hence the achievement scores announced to these *Ss* showed a parallel decline. Discrepancies between achievements and estimates (*D*-scores) tend to be slightly negative following success and slightly positive following failure. Changes in aspiration level lagged behind the changes in achievement. After exceeding their previous estimates, *Ss* raised their goals but not quite to the level of the last achievement. Similarly, after failing to reach their estimates, *Ss* lowered their goals but did not quite let them drop down to the level of their last achievement. These reactions to success and failure agree with those described in other investigations of the level of aspiration (8). The level of aspiration data indicate, then, that the experimental situation was associated with success for one of the experimental groups, and with failure for the other.

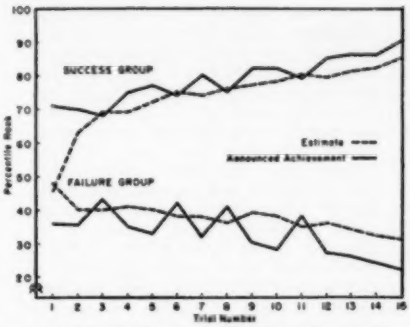

FIG. 1. SUBJECTS' ESTIMATES OF THEIR FUTURE PERFORMANCE AND ACHIEVEMENT SCORES ANNOUNCED TO THEM DURING THE EXPERIMENTAL TASK
Both estimates and achievements are in terms of percentile ranks.

Performance under success and failure. Further evidence for the differential reaction of the Success Group and the Failure Group comes from their performance in the task for which their level of aspiration was measured. It will be recalled that *Ss*' span of apprehension was tested by means of 15

slides, each of which had 12 items (capital letters, small letters, and numbers) on it. Table 2 shows, for the two experimental groups and the control group, the average number of responses given during this test

TABLE 2

AVERAGE NUMBER OF RESPONSES GIVEN IN SPAN OF APPREHENSION TEST BY DIFFERENT GROUPS AND PERCENTAGES OF CORRECT AND INCORRECT RESPONSES

GROUP	AV. NO. RESPONSES	AV. PER CENT CORRECT	AV. PER CENT INCORRECT
Failure	47.4	81.2	18.8
Success	44.7	87.8	12.2
Control	45.0	88.0	12.0

and the percentages of correct and incorrect responses. The members of the Failure Group gave a somewhat larger number of responses and made a larger percentage of errors than did the members of the Success Group and the Control Group. A graphic comparison of the error percentages is made

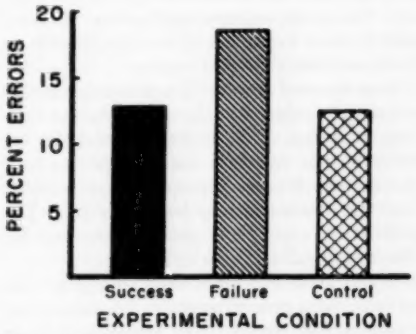

FIG. 2. PERCENTAGES OF INCORRECT RESPONSES GIVEN BY DIFFERENT GROUPS IN SPAN OF APPREHENSION TEST

in Fig. 2. Although the differences in the percentages of errors are not very large, they are consistent and statistically reliable. The mean difference in per cent errors between the Control Group and the Failure Group is significant ($t = 2.26$, $p < .01$) as is the difference between the Success Group and the Failure Group ($t = 1.96$, $p = .05$).

It appears that members of the Failure Group tried to counteract their failures by attempting to recognize a larger number of items than are within the normal span of apprehension. Being overmotivated and anxious, they made a larger number of mistakes. In a previous experiment investigating the effects of failure on perceptual efficiency Postman and Bruner (15) had found very similar results. After failing in an insoluble perceptual task, their subjects showed "recklessness" in their perceptual behavior, attempting premature interpretations of complex perceptual stimuli. As a result, their actual perceptual attainment suffered. Equally relevant is the experimental evidence that increase in motivation beyond an optimal point impairs the problem solving capacity of the organism (1). Perceiving, which in this situation may be described as instrumental behavior, is no exception to this generalization. The difference in the percentage of errors may thus be interpreted as evidence for effectiveness of the differential treatment of the two experimental groups.

To summarize: both the level of aspiration behavior of the *Ss* and the quality of their performance indicate that contexts of success and failure were created for the experimental groups. We are now ready to consider the resulting differences in sensitivity to words connoting success and failure.

Sensitivity to motivationally relevant words. Twelve motivationally relevant words (4 deprivation words, 4 instrumental words, and 4 goal words) were presented for recognition in the second part of the experiment. Each of the 12 critical words was matched with an equally familiar "neutral" control word (see Table 1). In order to obtain threshold measures independent of individual differences in tachistoscopic acuity *recognition thresholds for critical words were expressed as deviations from the thresholds for the matched control words.* Take, for example, the word *perfection* and its matched control word, *ingredient.* If it took a subject 10 trials to recognize *perfection* and 12 trials to recognize *ingredient,* the threshold for *perfection* would be −2; if he recognized *perfection* on the 12th trial and *ingredient* on the 10th trial, the threshold would be +2. Positive scores, therefore, indicate that the critical word is recognized more slowly than its matched control word; negative scores indicate speedier recognition of the critical word than of the

control word. Since individual differences in tachistoscopic acuity tend to be rather great, this transformation reduces the variability within groups and makes comparisons among groups more sensitive.

Table 3 and Fig. 3 show the average recognition thresholds for the three types of critical words under Success, Failure, and Control

TABLE 3

AVERAGE RECOGNITION THRESHOLDS FOR THE THREE TYPES OF CRITICAL WORDS UNDER THE VARIOUS EXPERIMENTAL CONDITIONS

Thresholds are expressed as differences between the number of trials required for recognition of critical and of control words.

GROUP	TYPE OF WORD		
	DEPRIVATION	GOAL	INSTRUMENTAL
Failure	+0.69	−0.12	−0.29
Success	+1.83	−1.50	−0.75
Control	+1.81	−0.60	−0.56

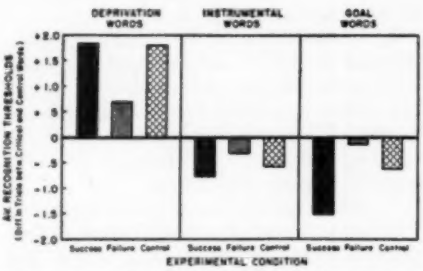

FIG. 3. AVERAGE RECOGNITION THRESHOLDS FOR THE THREE TYPES OF CRITICAL WORDS UNDER THE VARIOUS EXPERIMENTAL CONDITIONS

conditions. Let us consider the differences among the groups in their sensitivity to the three types of words. The tests of significance relevant to these comparisons are presented in Table 4.

Deprivation words. These words connote failure. Members of the Failure Group are more sensitive to these words than are members of the Success Group and the Control Group. The differences between the Failure Group and the other two groups are both significant. Control *Ss* and Success *Ss* have almost identical thresholds.

Goal words. These words connote success. Members of the Success Group have the lowest thresholds for this type of word; members of the Failure Group have the highest thresholds. The Control Group is in an intermediate position but is closer to the Failure Group than to the Success Group. The difference between the Success Group and the Failure Group is clearly significant; that between the Success Group and the Control Group approaches significance. Control and Failure *Ss* do not differ significantly in their sensitivity to goal words.

TABLE 4

TESTS OF SIGNIFICANCE OF DIFFERENCES AMONG THE VARIOUS GROUPS IN RECOGNITION THRESHOLDS FOR CRITICAL WORDS

GROUPS COMPARED	t	p
Deprivation Words *		
Failure-Control	1.93	.03
Failure-Success	2.00	.02
Success-Control	0.03	.99
Goal Words *		
Failure-Control	0.84	.20
Failure-Success	2.42	.01
Success-Control	1.32	.10
Instrumental Words **		
Failure-Control	0.46	.62
Failure-Success	0.81	.42
Success-Control	0.28	.76

* A one-tail test was used since the direction of the difference was predicted in the experimental hypothesis.
** A two-tail test was used since the direction of the difference was not predicted.

Instrumental words. These words connote striving, i.e., they are related to activities leading to success. The order of thresholds for instrumental words is the same as for goal words: the Success Group is most sensitive and the Failure Group least sensitive. The differences among the groups, however, fall short of significance in this case.

Apart from the differences among groups, is there not a general tendency for success and instrumental words to be recognized faster than failure words? The data plotted in Fig. 3 do not justify such a conclusion. We must remember that the recognition threshold for each critical word was expressed as a *deviation from the threshold of its control word.* Thresholds for the three classes of critical words were, therefore, not measured from a common base line, and comparisons among these classes cannot easily be made. We can, however, compare differences among groups of *Ss* for a given class of words since all *Ss* responded to the same critical and control words. Whether critical and control words were well matched in difficulty or not,

their relative difficulty was the same for all groups. Differences among groups can thus be legitimately attributed to the experimental treatment.

DISCUSSION

Experimentally created differences in situational context significantly influence recognition thresholds for words. Specifically, a context of success lowers thresholds for words connoting success; a context of failure lowers thresholds for words connoting failure. In terms of the theoretical scheme discussed above, hypotheses related to success gain strength—are readily aroused and confirmed—in a situation associated with success. By the same token, hypotheses related to failure gain strength in a situation associated with failure. This evidence suggests that situational cues may serve to arouse specific systems of hypotheses, thereby creating conditions that are favorable to the selective perception of objects to which the hypotheses refer. A strong hypothesis, once it has been aroused, will be readily confirmed in the presence of further appropriate stimulus cues.

Situational cues which will arouse a system of hypotheses may be many and varied. It may, indeed, be argued that through association any feature or combination of features of a situation may come to arouse a given type of hypothesis. Physical aspects of the situation, the presence of certain individuals, a particular temporal sequence of events—all of these may serve as situational cues arousing a particular set of hypotheses. We may speculate that some situational cues directly arouse perceptual hypotheses whereas in other cases it is some other response to the situational cues which mediates the arousal of perceptual hypotheses.[3] It is very probable that verbal labels and symbols which become attached to situations mediate the arousal of perceptual hypotheses in this manner. In this experiment, for example, *S*'s labeling of the situation as one of success or failure may well have mediated the differential arousal of hypotheses related to success and failure which then resulted in the observed threshold differences.

[3] See Miller and Dollard (12) for a development of the view that responses may function as cues. In this situation, verbal responses may function as cues for the arousal of perceptual hypotheses.

It is interesting to note in this connection that the pattern of thresholds for instrumental words is similar to that for success words but yields no significant differences among the groups. Examination of these words shows that two of them are very close to success words in their connotation (improve, achieve) and that none of them suggests failure. To the extent that instrumental words do indeed have success connotations, the conditions favorable to the recognition of success words are also favorable to the recognition of instrumental words. The influence of the experimental conditions on sensitivity to instrumental words may have been limited to this indirect effect. Such findings serve to remind us that the logical classification of stimuli by the experimenter does not necessarily correspond to the subject's predispositions to classify events (i.e., to his systems of hypotheses).

We have attributed the observed differences in perceptual selectivity to the arousal of specific sets of hypotheses in the presence of different situational cues. In general, perceptual selectivity varies with the degree of association between stimulus cues and predispositions to perceptual response (strength of hypotheses). The main point of our argument is, then, that the facts of perceptual selectivity can be derived from an application of the laws of associative learning to the establishment of perceptual hypotheses. We have already made reference to examples of such applications. Frequency and recency of past confirmation and degree of emphasis through motivational relevance are all conceived as determinants of hypothesis strength. To the extent that the laws governing the acquisition of hypotheses can be stated, the prediction of perceptual selectivity can be improved. Such is the case for motivationally relevant events no less than for "neutral" ones.

Our analysis provides an alternative to the invocation of such personality mechanisms as "repression" and "autistic tendencies" in explanation of differences in perceptual sensitivity. It seems to us that theoretical arguments relating perception to such general personality factors hold out little promise for the precise specification and experimental manipulation of the conditions under which variations in perceptual behavior occur. It

would be difficult, for example, to vary repressive tendencies experimentally in order to demonstrate their perceptual consequences.[4] Our formulation which relates perceptual selectivity to principles of learning is capable of experimental tests and will stand or fall with the results of such tests. Some verifications have already been made of the propositions concerning the determinants of hypothesis strength and their relation to perceptual sensitivity (4, 16, 18, 19).

SUMMARY

This experiment is concerned with the conditions governing perceptual selectivity to stimuli connoting success and failure. The hypothesis was tested that perceptual selectivity does not necessarily favor success-related (wish-fulfilling) over failure-related (wish-denying) symbols. Relative sensitivity to these two types of stimuli depends, at least in part, on the situational context in which they appear.

A level of aspiration situation was used to establish contexts of success and failure experimentally. The experimental task was a span of apprehension test. Members of the Success Group were allowed to exceed their stated goals on most of the trials; members of the Failure Group failed to reach their estimates on most of their trials. A Control Group performed the same task as the two experimental groups but did not make any estimates of its performance.

The level of aspiration of the Failure Group steadily declined, whereas that of the Success Group showed a progressive increase. In addition, the Failure Group attempted more responses and made a larger percentage of errors in the experimental task than did the Success Group or the Control Group. Thus, both level of aspiration behavior and performance in the experimental task indicate that contexts of success and failure were created for the experimental groups.

Twelve motivationally relevant words—4 "deprivation" words connoting failure, 4 "goal" words connoting success, and 4 "instrumental" words connoting instrumental striving—were presented for tachistoscopic recognition in the second part of the experiment. The Failure Group was significantly more sensitive to deprivation words than were the Success Group and the Control Group. Similarly, the Success Group was significantly sensitized to goal words. The pattern of thresholds for instrumental words was similar to that for goal words but the differences among groups fell short of statistical significance.

Experimental manipulation of situational context can, therefore, significantly influence thresholds for words connoting success and failure. The theoretical analysis of these results emphasizes the dependence of selective sensitivity on the association between situational cues and perceptual predispositions (hypotheses).

[4] A similar situation exists in the area of memory. Attempts to demonstrate repression in memory experiments have encountered extremely serious methodological difficulties.

REFERENCES

1. BIRCH, H. G. The role of motivational factors in insightful problem solving. *J. comp. Psychol.*, 1945, **38**, 295–317.
2. BRUNER, J. S. Personality dynamics and the process of perceiving. In R. R. Blake and G. V. Ramsey, (Eds.), *Perception: an approach to personality.* New York: Ronald Press, 1951.
3. BRUNER, J. S., & POSTMAN, L. Emotional selectivity in perception and reaction. *J. Pers.*, 1947, **16**, 69–77.
4. BRUNER, J. S., POSTMAN, L., & RODRIGUES, J. Expectation and the perception of color. *Amer. J. Psychol.*, 1951, **64**, 216–227.
5. FRENKEL-BRUNSWIK, E. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 1949, **18**, 108–143.
6. KLEIN, G. S., & SCHLESINGER, H. Where is the perceiver in perceptual theory? *J. Pers.*, 1949, **18**, 32–47.
7. LEVINE, R., CHEIN, I., & MURPHY, G. The relation of the intensity of a need to the amount of perceptual distortion, a preliminary report. *J. Psychol.*, 1942, **13**, 283–293.
8. LEWIN, K., *et al.* Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*, Vol. I. New York: Ronald Press, 1944.
9. MCCLELLAND, D. C., & ATKINSON, J. W. The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *J. Psychol.*, 1948, **25**, 205–222.
10. MCCLELLAND, D. C., & LIBERMAN, A. M. The effect of need for achievement on recognition of need-related words. *J. Pers.*, 1949, **18**, 236–251.
11. MCGINNIES, E. M. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
12. MILLER, N. E., & DOLLARD, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
13. MURPHY, G. *Personality.* New York: Harper, 1947.
14. POSTMAN, L. Toward a general theory of cognition. In J. H. Rohrer and M. Sherif (Eds.), *Social psychology at the crossroads.* New York: Harper, 1951.
15. POSTMAN, L., & BRUNER, J. S. Perception under stress. *Psychol. Rev.*, 1948, **55**, 314–323.

16. POSTMAN, L., & BRUNER, J. S. Hypothesis and the principle of closure: the effect of frequency and recency. *J. Psychol.*, 1952, 33, 113–125.

17. POSTMAN, L., BRUNER, J. S., & MCGINNIES, E. M. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, 43, 142–154.

18. POSTMAN, L., BRUNER, J. S., & WALK, R. D. The perception of error. *Brit. J. Psychol.*, 1951, 42, 1–10.

19. POSTMAN, L., & LEYTHAM, G. Perceptual sensitivity and ambivalence of stimuli. *J. Pers.*, 1951, 19, 390–405.

20. POSTMAN, L., & SCHNEIDER, B. Personal values, visual recognition, and recall. *Psychol. Rev.*, 1951, 58, 271–284.

21. SANFORD, R. N. The effect of abstinence from food upon imaginal processes. *J. Psychol.*, 1936, 2, 129–136.

22. THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Teachers College, Columbia Univ., 1941.

23. VANDERPLAS, J. M., & BLAKE, R. R. Selective sensitization in auditory perception. *J. Pers.*, 1949, 18, 252–266.

24. WOODWORTH, R. S. Reenforcement of perception. *Amer. J. Psychol.*, 1947, 60, 119–124.

Received April 9, 1951.

THE EFFECT OF AN AUDIENCE ON BEHAVIOR IN A CHOICE SITUATION [1]

SEYMOUR WAPNER AND THELMA G. ALPER
Clark University

THE one-way vision observation room has been widely adopted in recent years by both clinical and experimental psychologists for use in applied and research situations. In some cases the existence of the one-way mirror is not made known to the client or *S*. In other cases he is told either that he is being observed or that he could be observed. For the most part the assumption has been made that the one-way vision setup affects the behavior of the *S* very little, if at all.[2] This assumption, however, has not been subjected to adequate experimental analysis.[3] Indeed, analysis of this problem has wider implications than just the methodological one. For example, considered from the viewpoint of the effect of an audience on behavior, behavioral changes in the one-way vision setup are of theoretical interest to the clinician and to the social psychologist as well, since basically these changes bear on the problem of social interaction.

To study the effect of an audience on behavior, a situation is required which provides a sensitive, quantitative measure of meaningful behavioral changes. A choice situation fulfills these requirements and, in addition, has the advantage that the results would have implications for theories of decision.

The present experiment was designed to study the effect of an audience on the individual's behavior in a choice situation under different instructional sets using personality-oriented and neutral choice materials.

EXPERIMENTAL DESIGN AND PROCEDURE

A four-variable factorial analysis of variance design was used. The experimental variables were: (*a*) type of audience (varied three ways); (*b*) type of instructions (varied two ways); (*c*) sex of subject (varied two ways); and (*d*) experimenter (varied two ways). The design thus included twenty-four combinations of variables. The experiment was replicated five times: i.e., five subjects were assigned at random to each of the twenty-four combinations of variables, making a total of 120 *Ss*.

The details of the experimental design and the number of *Ss* in each cell are shown in Table 1.

Experimental Variables

Type of audience. The experiment was conducted in a one-way vision observation room. Three audience situations were used:

1. No audience. The curtains were drawn across the one-way vision mirror and only the *E* was present in the room with the *S*.

2. Unseen audience. The one-way vision mirror was exposed. *Ss* were informed of the nature of the mirror and were told that there were people behind it who could both hear and see what was going on in the experimental room.

3. Seen audience. The one-way vision mirror was exposed. The illumination was equalized between the experimental and observation rooms so that the audience could see the *S* and the *S* could see the audience.

In all conditions, the *E* remained in the experimental room with the *S* throughout the test. The unseen and the seen audiences behind the one-way vision mirror included a faculty member and three or four students. Both males and females were always included in the audience.

Type of instructions. Two instructions were used:

1. Task-oriented instructions. These were designed to emphasize that the *material* rather than the subject was being studied. This was accomplished by the following instructions: "I have some materials here which I would like to use in an experiment later on this term. Before I can use them, however, I need to know more about these materials. Some seem to be harder than others. I want to know which are the easy ones and which the really hard ones. After I've tried them out on a lot of people I can classify the materials more certainly as 'hard' or 'easy.'"

2. Ego-oriented instructions. These were designed to emphasize that the subject was being given a personality test. This was accomplished by the following instructions: "Have you ever taken any personality tests? We worked one out last year which tells us a great deal about an individual's personality. It consists of a set of sentences or phrases each of which is followed by two words.

[1] This study was reported at the annual meeting of the Eastern Psychological Association, Worcester, Massachusetts, 1950.

[2] See Allport (1, p.40).

[3] Experimentation on allied problems includes privately vs. publicly expressed opinion (2, 9, 10), the effect of spectators on individual efficiency (7), and closed vs. open balloting (4).

The word the person chooses gives us the clue to his personality. I am going to give you this test now."

3. Sex of subject. Sixty female and sixty male undergraduate *Ss* were used.

4. Experimenter. Two experimenters were used. A female *E* tested 30 of the female and 30 of the male *Ss*, and a male *E* tested the remaining half of the *Ss*.

Since only two *Es* were used, there is implicit in the design a confounding of experimenter and sex of experimenter. Differences involving the *E* variable could thus be a function of the differences between the *Es*, irrespective of sex, and/or of differences in the sex of the *Es*.

involving personality-oriented material (DP), e.g., the phrase "a bad habit" followed by the words "nailbiting-masturbation"; five easy choices involving personality-oriented material (EP), e.g., the phrase "a masculine characteristic," followed by the words "strong-weak"; fifteen difficult choices involving neutral material (DN), e.g., "a musical composition," followed by the words "symphony-concerto"; and five easy neutral choices (EN), e.g., "season of warm weather," followed by the words "winter-summer."

The phrase-word combinations used in the experiment were selected from a large sample of items pretested on a pilot group of 28 *Ss*. The usability of an item was determined on the basis

TABLE 1

EXPERIMENTAL DESIGN

AUDIENCE (A)	INSTRUCTIONS (I)	EXPERIMENTER (*E*)			
		E_1 SEX (S)		E_2 SEX (S)	
		MALE	FEMALE	MALE	FEMALE
None	Task-oriented	5	5	5	5
	Ego-oriented	5	5	5	5
Seen	Task-oriented	5	5	5	5
	Ego-oriented	5	5	5	5
Unseen	Task-oriented	5	5	5	5
	Ego-oriented	5	5	5	5

Materials

The experimental materials consisted of a series of phrases each of which was followed by two words.[4] The *S* was required to choose which of the two words fitted the given phrase more closely. Forty pretested phrase-word combinations ranging from easy to difficult discriminations and including an equal number of personality- and neutral-oriented items were used. Personality-oriented items were designed to emphasize characteristics or traits of people. Neutral-oriented items, however, had an impersonal reference, emphasizing concrete or abstract qualities of things. In the easy discriminations, no conflict of choice was involved since the alternative words had quite different meanings, and only one fitted the phrase. The difficult discriminations did involve conflict of choice since the two words were synonyms and both fitted the phrase. Ten of the forty phrase-word combinations were easy discriminations and thirty were difficult. Half of each type were personality-oriented and half were neutral-oriented. Thus, the forty phrase-word combinations included fifteen difficult choices

[4] Material of this general type was used in an experiment by Festinger and Wapner (8).

of percentage of choice by the pilot group of alternative words in a given phrase-word combination. An item was classified as difficult if one of the two alternatives was consistently chosen by not more than 75 per cent of the pilot group. In order to be included as an easy item one of the two alternatives had to be consistently chosen by 100 per cent of the pilot group.

The order of presentation of phrase-word combinations was established by means of a table of random numbers. The sequence thus established was used for all *Ss*. Approximately the same number of all types of items appeared in the first twenty and in the last twenty of the forty phrase-word combinations. Thus the materials in the two halves of the experimental session were comparable with respect to type and difficulty of item.

Apparatus

The apparatus had two main parts, an exposure apparatus and a timing device. The front face of the exposure apparatus had three open slots, one to accommodate the phrase and one for each of the two alternative words. The phrase slot was located above the slots for alternative words and was exposed to the *S*'s view at all times. The other two

slots were covered by a shutter which was manually controlled by *E*. The handle which released the shutter had a mercury switch in it which activated the timing device. The timing device consisted of a standard electric clock graduated in .01 seconds, and an electronic voice key. The clock started when *E* released the shutter. The *S*'s verbal response into the microphone stopped the clock.

The phrase-word combinations were printed in large type on firm white cloth attached to rollers behind the apparatus. By winding the rollers *E* changed the phrase-word combination for succeeding trials.

Procedure

The *S* was brought into the testing room and seated in front of the exposure apparatus. Directions appropriate to the condition, i.e., type of audience and type of instructions, were given verbally. The procedure was then demonstrated to *S* by means of two sample phrase-word combinations. No mention was made that *S* was being timed.

As soon as *S* had read a given phrase he was instructed to signal this fact by raising his right hand. At this signal, *E* said "Ready" and released the shutter, thus exposing the alternative words. *S* chose the alternative which seemed to him to fit the phrase more closely, verbalizing this choice into the microphone. The time to make a choice (decision-time), accurate to .01 second, was the measure analyzed.

RESULTS

Types of Items

As noted under materials, the four types of items, (DP, DN, EP, EN) constructed for use in the experiment, differed with respect to orientation and difficulty. In designing the personality- and neutral-oriented items no criteria were used other than the judgment of the *E*s. The distinction between difficult and easy items, however, was made on the basis of an independent criterion, namely, the percentage of choice by a pilot group for one of the two words in a phrase-word combination. Whether types of items so selected will yield decision-time differences requires statistical analysis.

Differences among the four types of items were evaluated by a two-way classification analysis of variance with subjects as one variable and type of material as the other. The results of the analysis are presented in Table 2. The *F*-test for types of material is statistically significant beyond the 1 per cent level. These findings indicate that there are significant differences among the four types of material. The mean decision-time in milliseconds is longest for DP (208 ms.), next for DN (187 ms.), next for EP (154 ms.), and shortest for EN (137 ms.) items.

TABLE 2

DIFFERENCES IN DECISION-TIME FOR THE FOUR TYPES OF CHOICE MATERIAL

SOURCE OF VARIATION	*df*	SUM OF SQUARES	VARIANCE ESTIMATE	*F* *
Types of Material (M)	3	367,341.5	122,447.2	30.19 **
Subjects (S)	119	2,838,675.5	23,854.4	5.88 **
Residual (M×S)	357	1,447,981.0	4,056.0	
Total	479	4,653,998.0		

* Residual variance was used as denominator.
** Significant at less than 1 per cent level.

To determine where the differences lie, *t*-tests were done. Each type of item differs significantly from every other type of item. All differences are significant at less than the 1 per cent level, with the exception that the difference between EP and EN is at or below the 5 per cent level.

It can also be noted from Table 2 that the *F*-test for subjects is significant. Since the analysis does not evaluate individual differences within an experimental condition this significant *F*-test is a function of subject and/or differences between experimental conditions. Additional analyses of variance, done for each of the twenty-four combinations of variables, showed significant differences between subjects. This finding is consistent with the results of other investigations (cf. 8) which have demonstrated stable individual differences in decision-time.

Experimental Variables

Because clinicians suggest that in the course of time individuals adjust to being observed, it seemed advisable to analyze the data separately for the first and second half of the experimental sessions. It is possible that an experimental condition (for example, audience) may differentially affect decision-time in the early part of the experiment; yet, in the later part such effects may be so dissipated that the audience variable is no longer significant.

Analyses of variance for each of the four types of material, done separately for first and second half of the experimental session, are presented in Table 3.

TABLE 3

RESULTS OF EIGHT ANALYSES OF VARIANCE: SEPARATE FOUR-VARIABLE ANALYSIS FOR FIRST AND SECOND HALF OF EXPERIMENTAL SESSION FOR EACH OF THE FOUR TYPES OF MATERIAL

SOURCE OF VARIANCE		DIFFICULT CHOICE-PERSONALITY				DIFFICULT CHOICE-NEUTRAL				EASY CHOICE-PERSONALITY				EASY CHOICE-NEUTRAL			
		1ST HALF		2ND HALF		1ST HALF		2ND HALF		1ST HALF		2ND HALF		1ST HALF		2ND HALF	
	df	VARIANCE ESTIMATE	*F* ***	VARIANCE ESTIMATE	*F*	VARIANCE ESTIMATE	*F*	VARIANCE ESTIMATE	*F*	VARIANCE ESTIMATE	*F*	VARIANCE ESTIMATE	*F*	VARIANCE ESTIMATE	*F*	VARIANCE ESTIMATE	*F*
Audience (A)	2	49,618.1	5.02**	23,615.7		58,007.8	4.24*	30,957.8		14,541.5		14,189.1		13,619.8	12.13**	8,413.4	
Instruction (I)	1	35,327.6		59,853.3		69,600.8	5.09*	36,540.3		14,213.1		5,617.0		72.1		1,171.9	
Sex of subject (S)	1	30,431.5		2,484.3		16,240.1		3,221.9		11,097.2		444.7		85.0		1,435.2	
Experimenter (E)	1	7,441.6		25,111.4		12,689.6		10,010.1		488.0		2,585.4		72.1		46.9	
A×I	2	7,124.7		17,276.0		18,212.6		12,558.9		12,824.6		14,986.0		1,361.3		6,777.8	
A×S	2	1,982.4		2,994.2		3,042.8		290.2		16,527.5		18,189.4		1,854.8		5,171.4	
A×E	2	2,132.7		9,099.0		12,332.6		3,911.5		13,843.0		6,781.8		108.2		1,258.0	
I×S	1	47,161.8	4.78*	19,763.4		42,638.7		6,600.1		9,257.3		3,887.4		2,950.5		1,274.0	
I×E	1	74,448.0	7.54**	43,095.1		75,000.0	5.48*	12,648.6		2,447.9		658.0		785.4		156.4	
S×E	1	33,632.7		44,239.5		55,900.8	4.08*	34,734.2		14,213.1		15,030.4		696.0		14,807.4	
A×I×S	2	1,735.8		1,847.1		4,249.8		3,430.1		12,068.3		11,754.4		5,382.3	4.80*	7,194.0	
A×I×E	2	5,797.8		1,894.7		2,908.6		651.0		10,445.8		5,287.9		1,859.8		161.4	
A×S×E	2	7,158.9		30,223.2		25,284.0		28,738.9		11,215.2		10,498.1		159.8		6,509.6	
I×S×E	1	47,400.0	4.80*	31,425.2		40,922.1		19,381.3		9,540.4		7,792.4		1,039.9		5,562.4	
A×I×S×E	2	3,268.7		31,284.7		5,176.9		20,332.1		10,009.6		9,447.8		2,315.4		3,108.8	
Within	96	9,874.2		16,225.3		13,684.0		12,966.3		12,945.5		9,585.8		1,122.5		4,679.5	
Total	119																

* Significant at or below 5 per cent level of confidence.
** Significant at or below 1 per cent level of confidence.
*** The denominator of the *F*-ratio was the within groups variance. Only significant *F*-tests are entered in the table.

Audience. The *F*-value for the audience variable in the first half of the experimental session is significant at or below the 5 per cent level for all types of material other than EP. For all types of material decision-time is shortest for no audience, intermediate for seen audience, and longest for unseen audience, as shown in Table 4. In those instances

TABLE 4

MEAN TIME (MILLISECONDS) TO MAKE A CHOICE FOR AUDIENCE CONDITIONS IN THE FIRST HALF OF THE EXPERIMENTAL SESSION

AUDIENCE	TYPE OF MATERIAL			
	DP	DN	EP	EN
None	179	178	137	117
Seen	204	207	168	119
Unseen	248	254	171	150

where the *F*-value is significant, *t*-tests reveal that the mean for no audience is significantly different from the mean for unseen audience; no audience is not significantly different from the seen audience; and, finally, the seen audience is significantly different from the unseen audience only for DP and EN types of material. In general, then, the unseen audience yields significantly longer decision-times than the other two audience conditions, none and seen audience. The latter two conditions, however, do not differ significantly from each other.

In contrast to the significant audience effects for the first half of the experimental session, there is no evidence of significant audience effects for the second half.

Instructions. In only one instance, DN first half, was the instruction variable significant. Here, decision-time was shorter for ego-oriented instructions than for task-oriented instructions.

A number of significant interaction effects occur involving the instructions variable. These effects are found only in the first half of the experimental session. For DP there are three significant interactions: instructions by sex of subject (I×S), instructions by experimenter (I×E), and a significant triple interaction, instructions by sex of subject by experimenter (I×S×E). Inspection of the I×S interaction reveals that task-oriented instructions yield longer decision-times for the male experimenter while ego-instructions yield longer decision-times for the female experimenter. The I×S×E interaction shows that for the female experimenter, ego-instructions yield the longer decision-times for both male and female *S*s. In contrast, for the male experimenter task-oriented instructions yield the longer decision-times for both male and female *S*s.[5]

For the DN type of item a significant I×E interaction occurs in the first half of the session. The direction of the decision-times parallels that described above for the I×E interaction with DP items: namely, task-oriented instructions yield longer decision-times for the male experimenter. It is interesting to note at this point that differences between *E*s occurred with respect to the instructions-variable even though considerable time was spent in standardizing the instructions prior to running the experiment proper.

Sex of subject. Sex differences as such yield no significant decision-time differences. However, a significant interaction involving sex of subject and experimenter (S×E) did occur for one type of item, viz., DN items for the first half of the experimental session. Here it was found that decision-time is longer for male *S*s when the *E* is female and longer for female *S*s when the *E* is male. With respect to this interaction it is important to note, as indicated above, that in the experimental design employed here there is a confounding of *E* and the sex of *E*. This means that differences with respect to the experimenter-variable may be a function of experimenter and/or sex of experimenter.

Experimenter. Experimenter differences as such yield no significant decision-time differences. The only significant differences involving the *E* variable are the interactions already noted.

Adjustment to Experimental Conditions

It should be emphasized that the experimental variables yield significant differences only in the first half of the experimental session. This finding has important bearing

[5] When the triple interaction is used as the error term to evaluate the double interactions, the latter are not significant.

on how *Ss* adjust to this experimental situation. Apparently, even within a twenty-minute test period, adjustment to the experimental situation takes place: the significant differential effects for the audience variable in the first half of the experimental session are dissipated during the second half of the experimental session so that the audience variable is no longer significant.

The nature of the adjustment to the experimental situation can be further analyzed by determining whether there is an increase or a decrease of decision-time from first to second half of the session. For this purpose the mean decision-time of each *S* for the twenty items in the first half of the experimental session was compared with his mean decision-time for the twenty items in the second half. It was found that for all *Ss* combined, regardless of experimental condition, 72 per cent showed a decrease in decision-time from first to second half of the experimental session. This percentage, tested against the hypothesis that a fifty-fifty distribution of increases and decreases in decision-time would occur by chance, yielded a chi-square of 22.54, which is significant beyond the .01 level of confidence. In general, then, decision-time becomes shorter in the second half of the experimental session.

The possibility that this decrease in decision-time is differentially affected by the type of audience condition was investigated by tabulating the number of increases and decreases within each of the three audience conditions. A chi-square test of independence shows no evidence of a significant relationship between audience condition and change in decision-time from first to second half of the experimental session. However, it is interesting to note that more *Ss* show a decrease in decision-time in the unseen audience than in the other two conditions.

DISCUSSION

The main finding in the present study is that the presence of an audience affects time to make a choice. Of the three audience conditions, time to make a choice is longest in the presence of an unseen audience, shortest when no audience is present, and intermediate when the audience is seen by the *S*. Increase in time to make a choice is taken as evidence that *S* experiences greater conflict in making the choice.

It is well-known that difficulty of choice (the more similar the alternatives the greater the difficulty in choosing between them) affects the degree of conflict experienced by the *S*.[6] This is attested to in the present experiment by the significant differences obtained between easy and difficult choices for both personality and neutral material. Here the degree of similarity of alternatives affects the degree of conflict experienced by the *S*. The question arises, then, as to how factors in the situation other than the similarity of the alternatives can influence the conflict experienced by the *S* in making a choice.

An answer to this question is suggested by the significant difference obtained between the difficult-personality and difficult-neutral material. Here the increase in conflict of choice can be considered a function of the relevance of the material to the self. Similarly, the presence of an audience has relevance for the self. For example, an audience may serve to threaten self-status (need to be thought well of by others). The audience, after all, is a potential interpreter of the choices made by the individual. An audience which threatens self-status should serve as a greater restraining force[7] in making a decision than an audience interpreted as non-threatening. The significantly increased decision-time in the presence of an unseen audience is understandable from this viewpoint: an audience who cannot be seen but is "out there" watching and listening to the choices being made is indeed more threatening to self-status than an audience whose composition is known.

The sex of subject by experimenter interaction can be interpreted along similar lines. Although it cannot be ascertained from the present experimental design whether sex of experimenter per se is the essential variable, the interaction obtained suggests that the *composition* of the audience interpreting the choices is probably a crucial variable.

The hypothesis that ego-oriented instructions would yield longer decision-times than task-oriented instructions was not supported.

[6] See, for example, Barker (3) and Cartwright (5).

[7] The term "restraining force" refers to the force which prevents the subject from making a decision. (See Cartwright and Festinger [6]).

In one instance (difficult-neutral materials) decision-time was significantly shorter for the ego-oriented than for the task-oriented instructions. This unexpected finding may be due to the faulty wording of the task-oriented instructions. These instructions, in stressing that the *E*s were concerned to determine which of the materials were easy, and which difficult, put a premium on correctness of choice. Hindsight now makes it clear that under such task-oriented instructions, a spurious variable, correctness of choice, is entering into the decision-time. If this explanation is correct, the importance of the wording of the instructions to establish a given set cannot be overstressed.

The fact that all significant findings are confined to the first half of the experimental session suggests that the significant variables have only ephemeral effects. The restraining forces are effective during the first half of the experimental session but as the session continues they lose their potency, i.e., adjustment takes place. It is possible that if the restraining forces are made more potent, for example, if more heavily loaded choices such as, "the person you love best: mother-father," were used throughout the series, adjustment might not take place during the experimental hour.

The assumption that adjustment takes place rapidly is commonly made by clinicians and experimenters alike who use one-way vision screens. Such an assumption is warranted within the context of the results of this experiment. However, the question of how restraining forces *more* potent than those used in the present experiment affect adjustment is still a matter for research.

The results of this experiment have bearing on the theory of decision developed by Cartwright and Festinger (6). Their theory has two main concepts: (*a*) driving forces and (*b*) restraining forces. Driving forces in a choice situation function in the direction of making a decision. There is a driving force which acts upon the *S* in the direction of each alternative. The resultant of the driving force (difference in strength between the driving force for one alternative compared to the others) determines the direction of the decision and the potency of the force to make a decision. Restraining forces prevent the *S* from making a decision. Cartwright and Festinger mention that such an attitudinal factor as "need to make a correct choice" serves as a restraining force. The present experiment has identified some of the variables which can serve as restraining forces. These variables include nature of the material and the audience.

Insofar as the experiment has shown that the individual's behavior in a choice situation is affected by the presence of an audience, the results have implications for the larger problem of social interaction.

SUMMARY

1. The effect of audience, of instructions and of materials on time to make a choice was studied.

2. Time to make a choice was longest in the presence of an unseen audience, next in the presence of a seen audience, and shortest when there was no audience other than the experimenter. The significant differential effects of the audience variable occur for the first half of the experimental session only.

3. Significant differences in time to make a choice were obtained for material which varied with respect to difficulty of choice, and to reference (personality vs. neutral). Items which involved difficult choices yielded longer times than easy choice items. Items with personality reference yielded longer times than neutral items.

4. Contrary to expectation there was some indication that time to make a choice is longer for task-oriented than for ego-oriented instructions.

5. Theoretical implications of the results are discussed.

REFERENCES

1. ALLPORT, G. W. "How shall we evaluate teaching?" In B. B. Cronkhite (Ed.), *A Handbook for College Teachers.* Cambridge: Harvard Univ. Press, 1950.
2. ALPER, THELMA G., & SOLOMON, R. L. Privately vs. publicly expressed attitudes toward rights for women. (Unpublished study.)
3. BARKER, R. B. An experimental study of the resolution of conflict by children. In *Studies in Personality, Contributed in Honor of Lewis M. Terman.* New York: McGraw-Hill, 1942.
4. CANTRIL, H. *Gauging public opinion.* Princeton: Princeton University Press, 1944.
5. CARTWRIGHT, D. Decision-time in relation to the differentiation of the phenomenal field. *Psychol. Rev.*, 1941, 48, 425–442.

6. CARTWRIGHT, D., & FESTINGER, L. A quantitative theory of decision. *Psychol. Rev.*, 1943, **50**, 595–621.
7. DASHIELL, J. F. An experimental analysis of some group effects. *J. abnorm. soc. Psychol.*, 1930, **25**, 190–199.
8. FESTINGER, L., & WAPNER, S. A test of decision-time: reliability and generality. (CAA Div. Res. Rep. No. 48, 1945; Publ. Bd. No. 50320.) Washington, D. C.: U. S. Dept. Commerce, 1947.
9. SCHANCK, R. L. A study of a community and its groups and institutions conceived of as behaviors of individuals. *Psychol. Monogr.*, 1932, **43** (2), Whole No. 195.
10. WILLIAMS, F., & CANTRIL, H. The use of interviewer rapport as a method of detecting differences between 'public' and 'private' opinion. *J. soc. Psychol.*, 1945, **22**, 171–175.

Received May 3, 1951.

DEFENSE AGAINST EGO-THREAT IN MEMORY AND PERCEPTION

CHARLES W. ERIKSEN
The Johns Hopkins University

THE occurrence of perceptual defense has been well established (2, 3, 4, 6, 7). Stimuli which are ego-threatening owing to association with unacceptable needs have higher recognition thresholds than neutral stimuli. The experiments demonstrating this have used a correlational approach, showing relationships between the degree of ego-threat and recognition threshold. Although significant, these correlations have not been large. The size of the unexplained variance suggests that variables other than degree of ego-threat may be important in the occurrence of perceptual defense.

Clinical experience amply demonstrates that people respond to ego-threat in a variety of ways. Accordingly it has been suggested that type of defense or defense mechanism may be a factor in the perception of ego-threatening stimuli (3). In support of this, Lazarus and Eriksen (6) found that neurotic patients who used intellectualizing defense mechanisms had lower thresholds for threatening material than did patients using repression and avoidance.

A large variety of ego defense mechanisms are recognized. An important way in which they differ is in how they affect the individual's behavior in the presence of the threatening stimulus. Mechanisms such as denial and repression produce avoidance, an attempt to keep the stimulus from awareness. Other mechanisms may permit a ready recognition of the threatening stimulus but self-esteem is preserved by rationalizing or explaining away the threat, or perhaps projecting it outward. In one case the emphasis of the defense is on the denial of external reality while in the other case the denial is in terms of how this reality applies to the individual. If this analysis is correct, we can see how type of ego defense would be an important variable in the perception of ego-threat.

The present experiment shows that prior knowledge of the effect of ego-threat upon memory enables one to predict with considerable accuracy the individual's perceptual response to threat. Individuals who have an over-readiness to forget (avoid) ego-threatening experiences show more defense at the perceptual level than do individuals whose response to such threat is more realistic.

The hypothesis for the present study may be stated as follows: In a situation where failure is ego-threatening, individuals who have a high success-to-failure recall ratio will show a greater degree of perceptual defense than will individuals with a low success-to-failure recall ratio.

EXPERIMENTAL DESIGN

The over-all plan involved two subexperiments. Experiment I was a memory for uncompleted tasks. An experimental group was given a series of tasks so designed that they would be successful on half and fail the other half. The tasks were administered under the guise of an intelligence test. To increase the deception stooges were present who feigned successful completion of all of the tasks.

A control group of subjects was administered the tasks under identical conditions with the exception of stooges and the belief that the tasks measured intelligence.

Experiment II determined the relation between association time for a word and its perceptual recognition threshold. The subjects were those in the experimental group in Experiment I who had a very high or a very low success-to-failure recall ratio.

PROCEDURE

Experiment I

The subjects in this experiment were 61 male students enrolled in introductory courses in psychology at Johns Hopkins. Part of the course requirements were that students serve a minimum of five hours as experimental subjects. The experimental group contained 39 *Ss* and the control group 22.[1]

[1] The experimental and control groups originally contained 54 and 30 subjects, respectively. Fifteen subjects were dropped from the experimentals and 8 from the controls owing to their failure to complete the solvable sentences.

The tasks were 14 sentences divided into two- and three-word phrases. The phrases for each sentence were presented in a jumbled order and it was *Ss* task to rearrange the phrases so as to obtain a meaningful sentence. The sentences were selected from among those used by Alper (1). Seven of the sentences were those classed by Alper as easily solved and the remaining seven were unsolvable.

To permit group administration the sentences were presented in a booklet with one jumbled sentence to a page and a blank page separating the individual sentences. The *Ss* were cautioned against reviewing or looking ahead. Space was provided on the page with the sentence for *S* to write out his solution. Two and a half minutes were given for the solution of each of the sentences.

To facilitate recall, each sentence was given a name by selecting a two-word phrase from the sentence, underlining this phrase and placing it above the other phrases. In addition, *E* called out the name of each sentence when the group began working on it.

The sentences were administered to the experimental group under the pretext that it was an intelligence test. They were told that *E* was engaged in research upon abstract thinking and concept formation and he needed *Ss* for this research. The intelligence test was being given as a means of selecting students who were intelligent enough to participate in the abstraction experiment. The *Ss* were further told that anyone with an IQ of 115 or better should be able to complete all or practically all of the sentences within the time limit.

To increase the experimental group's feeling of failure, stooges were scattered throughout the group. These stooges were classmates who were wise to the experiment. They pretended successfully to complete all the sentences.

The control group was administered the sentences under the impression that this was some material that *E* had dreamed up to keep *Ss* from rehearsing during a rest period in a learning experiment. They were told that they were to aid *E* in selecting sentences which would keep *Ss* in the learning experiment occupied for the necessary length of time. The control group was also told that many of the sentences could not be solved within the time provided. Outside of this difference in set and the absence of stooges, the experimental and control groups were treated identically.

After the administration of the sentences both groups spent five minutes writing down all the four-letter words they could think of beginning with the letter *t*. A five-minute recall period followed.

Experiment II

From *Ss* in the experimental group who recall four or more sentences in Experiment I, the seven who had the highest and the seven who had the lowest difference score of failures minus successes recalled were used as the subjects in Experiment II.

A word-association list containing 93 words was individually administered to these *Ss*. The list was a slight modification of the one used by Bruner and Postman (2). It consisted almost entirely of five-letter words. Association time was measured in hundredths of a second by an electric timer.

A group of 25 words was selected for each *S*. The five words to which *S* had the longest and the five words to which he had the shortest association time were chosen. Five words with medium long, five with medium, and five with medium short association times made up the additional 15 words. Approximately two weeks after the association test each *S*'s recognition threshold was individually determined for his list of words.

To determine recognition thresholds, each word was typed with an electric typewriter, photographed, and made into a 35 mm slide. The words were individually exposed on a ground-glass screen by means of a projector with a shutter attachment. A constant exposure of 10 milliseconds was used. Thresholds were obtained by varying the intensity of illumination through the use of a variac in the projector circuit. Voltage was varied from 30 to 50 volts, an illumination range of .35 to 2.61 millilamberts.[2]

The *S*'s thresholds were obtained by randomly dividing the 25 words that had been selected for him into two groups of 12 and 13 words each. All the words in a group were exposed at each illumination level before proceeding to the next higher level. Between each illumination level the slides were shuffled so that neither *S* nor *E* knew the order in which the words would be exposed. After all the words in Group I had been correctly recognized, the same procedure was followed for Group II. A criterion of three successive correct recognitions was used.

The *Ss* were instructed to guess. They were told that the words to be exposed were of all types. Some might be swear words or "dirty" words so they should not hesitate to report any guesses they might have.[3]

Association time and recognition scores were obtained by the use of intra-individual statistics. Each *S*'s association times for the 93-word list were normalized via a reciprocal transformation, and the mean and SD computed. The *S*'s association score for any word was a standard score based upon his own mean and variability.

An *S*'s recognition score for a word was computed in the following manner. His mean recognition threshold in volts was separately determined for his 12 words in Group I and 13 words in Group II. His *recognition score* for any given word was the deviation of its recognition threshold from the mean of the group in which it was a member.

RESULTS

Experiment I

In Table 1 the experimental and control groups are compared on the total number

[2] Within this range a unit change in voltage is linearly related to log millilamberts.

[3] Actually only four of the words used might be considered to come under this classification. They were penis, whore, teats, and raped.

of sentences recalled. The control group recalls slightly more, but the mean difference is so small as to be well within chance expectancy. In terms of variability, both groups are identical.

TABLE 1

COMPARISON OF EXPERIMENTAL AND CONTROL GROUPS ON TOTAL SENTENCES RECALLED

	Experimental ($N = 39$)	Control ($N = 22$)	t	F	p
Mean	5.4	5.7	.58		>.50
S^2	4.3	4.3		1.0	>.20

Our major interest, however, is in group differences in the recall of completed and incompleted sentences. To make this analysis a difference score was computed for each *S*. It was obtained by subtracting the number of completed sentences recalled from the incompleted sentences recalled. In Table 2 the experimental and control groups are compared on these difference scores. The experimental group recalls more completed (successful) sentences while the controls recall more incompleted. The mean difference, although not large, is significant beyond the .05 level.

TABLE 2

COMPARISON OF EXPERIMENTAL AND CONTROL GROUPS ON INCOMPLETED MINUS COMPLETED SENTENCES RECALLED

	Experimental ($N = 39$)	Control ($N = 22$)	t	F	p*
Range	+5 to −5	+3 to −2			
Mean	−.86	+.41	1.87		<.05
S^2	14.13	2.18		6.47	<.001

* The probabilities are for a one-tailed test.

The most striking difference between the experimental and control groups is in the variability of the difference scores. The variance of the experimental group is more than six times greater than that of the control. This greater variance appears to reflect individual differences in response to ego-threat. Some *Ss* respond by dwelling upon the threat (failure) whereas others respond by emphasizing their successes and forgetting their failures.

Experiment II

Here we are concerned with differences in the perceptual recognition of ego-threat between those *Ss* who recall failures (Failure-Recall Group) and those who recall successes (Success-Recall Group).

Table 3 gives the correlations obtained between association and recognition scores for each *S* in the Success-Recall and the Failure-Recall groups. The correlations for all *Ss* in the Success-Recall group are positive and are significant at or beyond the .01 level. For this group ease of perceptual recognition is inversely related to degree of ego-threat as measured by association time.

On the other hand, the correlations obtained for the Failure-Recall group are considerably lower. Only three are significant at the .05 level and one of these is negative. All *Ss* in this group show a lower correlation between association and recognition than does any *S* in the Success-Recall group.

TABLE 3

CORRELATIONS BETWEEN ASSOCIATION TIME AND RECOGNITION THRESHOLDS FOR THE SUCCESS-RECALL AND FAILURE-RECALL GROUPS

Subject	Success-Recall Product-Moment Correlation	z	Subject	Failure-Recall Product-Moment Correlation	z
A	.76	1.00	a	.37	.39
B	.74	.95	b	.37	.39
C	.64	.76	c	.31	.32
D	.60	.69	d	.27	.28
E	.58	.66	e	−.01	−.01
F	.56	.63	f	−.22	−.22
G	.50	.55	g	−.35	−.37
Mean z score		.75			.11
Combined variance estimate				.062	
		$t = 4.77$		$p < .001$	

N equals 25 for each subject. For a one-tailed test z values of .36 and .50 are significant at the .05 and the .01 levels.

We can determine whether the average of the correlations obtained for the Success-Recall group is reliably higher than the average obtained for the Failure-Recall group. By transforming the individual correlations into Fisher's z we can compare the mean z scores for the two groups. As you again see from Table 3, this comparison shows a highly significant difference. The Success-Recall group does have a greater relationship between degree of ego-threat and perceptual recognition.

According to our hypothesis the difference in perceptual behavior between the Success-Recall and Failure-Recall groups is due to their different ways of responding to ego-threat. Before making this interpretation, however, we must consider certain other possibilities.

The difference in correlations between the two groups is *not* due to differences in performance on the word association test. This you can see from Table 4. Here the groups are compared on the average of mean association times and average of the SD's on word associations. Differences between the two groups are well within chance expectancy.

TABLE 4

COMPARISON OF SUCCESS-RECALL AND FAILURE-RECALL GROUPS ON THE MEAN OF THEIR ASSOCIATION TIMES AND VARIABILITY

	SUCCESS-RECALL ($N=7$)	FAILURE-RECALL ($N=7$)	t
Average of mean association time	.00648	.00614	.67
Average of SD's of association time	.00164	.00183	.98

* All association times were normalized via a reciprocal transformation. The means and standard deviations in the table are in terms of these reciprocal values.

Another possibility is that the shape or structure of a word may influence its ease of recognition. Words were individually chosen for each *S*. Hence, there is the slight possibility that the differences between the Success-Recall and Failure-Recall *S*s is the result of some kind of bias in the selection of their words. To show that this was not the case, the following analysis was used.

An average recognition value was determined for each of the 76 words used among the 14 subjects. This value was the average of all the recognition scores (based upon both groups) that had been obtained for the particular word (see Procedure for definition of recognition score). Each *S*'s recognition score for a word was expressed as a deviation from the total group average. This deviation was then correlated with his association time score for the word. After transforming the correlations into Fisher's z, the mean z scores for the Success-Recall and Failure-Recall groups were compared. Table 5 summarizes this comparison.

TABLE 5

CORRELATIONS BETWEEN ASSOCIATION TIMES AND RECOGNITION THRESHOLDS FOR THE SUCCESS-RECALL AND FAILURE-RECALL GROUPS WHERE RECOGNITION THRESHOLDS HAVE BEEN CORRECTED FOR TOTAL GROUP PERFORMANCE

SUCCESS-RECALL			FAILURE-RECALL		
SUBJECTS	r	z	SUBJECTS	r	z
A	.49	.54	a	.02	.02
B	.46	.50	b	.11	.11
C	.24	.25	c	.05	.05
D	.29	.30	d	.21	.21
E	.43	.46	e	.04	.04
F	.30	.31	f	.00	.00
G	.29	.30	g	−.13	−.13
Mean z score:		.38			.04
Combined variance estimate:		.012			
		$t=5.73$		$p<.001$	

The Success-Recall group still shows a significantly higher mean correlation between association time and perceptual recognition. The differences in perceptual recognition behavior between the Success- and Failure-Recall groups cannot be attributed to a biased selection of words for one of the groups.[4]

The frequency with which a word occurs in ordinary usage has been reported to be an important factor in determining its recognition threshold (5). In the present study, however, frequency of occurrence plays only a minor role in determining either associa-

[4] The reduction in the size of the correlations is an artifact of the present comparison. The average recognition value for each word is based upon a small number of subjects, all of whom are in the sample. Of the 76 words used, the average recognition value of 49 is computed from three or less subjects. Thus the result is to minimize the deviation of any one of these subject's recognition scores.

This artifact does not invalidate the above group comparison. Rather it acts to increase the severity of the test.

tion time or recognition. This you can see from Table 6. Here the average association score for a word and its average recognition value have both been correlated with the log of the word's frequency in the combined Thorndike-Lorge count (8). Although significant, these correlations are quite low.

TABLE 6

CORRELATION OF ASSOCIATION TIME AND RECOGNITION THRESHOLD WITH THORNDIKE-LORGE FREQUENCY

	N	*r*	*p*
Thorndike-Lorge frequency vs. average association score	93	.30	<.01
Thorndike-Lorge frequency vs. average recognition value	76	.26	<.02

The significance of *r* is evaluated via Fisher's *z* transformation. The probabilities are for a one-tailed test.

DISCUSSION

The above results clearly indicate that the way an individual responds to ego-threat is an important variable in need and perception research. Those who respond to failure-induced ego-threat by forgetting their failures show a similar defense in perception. In both cases their response to ego-threat is to avoid its recognition, to keep it from awareness. In memory this defense takes the form of poorer recall for the threatening stimulus whereas in perception it takes the form of failure to recognize or perceive.

On the other hand, individuals whose memory for ego-threat is not impaired tend to perceive ego-threatening stimuli as readily as they do nonthreatening stimuli. Their performance in both situations is more reality-oriented, which is to say that their objective performance is less distorted or influenced by their own needs.

It is interesting to speculate that the defensive mechanisms of these individuals are equivalent to rationalizing or intellectualizing mechanisms which explain away the threat and thus permit ready recognition of the stimulus.

In the same vein it is tempting for us to equate the defensive behavior of the Success-Recall group with the ego mechanisms of repression and denial. But for both groups such conclusions are premature. Considerably more experimental data are needed on ego defenses before we can even accept our general classification of defenses as adequate.

The present results have several important implications for personality research. In addition to showing the role of ego defense in perception, the present data effectively answer a major criticism that has been leveled against need and perception research in general. Howes and Solomon (5) have argued that the differential thresholds obtained for need-related words is primarily a function of the subject's familiarity with the stimulus. Thus they have emphasized the correlation between the recognition threshold for a word and its frequency in the Thorndike-Lorge word counts.

A general criticism of their position has been given elsewhere (6). Here, however, we should note that the present data cannot be explained on a simple familiarity hypothesis. Whether or not a person tends to recall his successes or his failures has no relation to his familiarity with words. Yet on the basis of such a distinction we find that we can effectively predict an individual's correlation between association time and recognition threshold. That familiarity can and does play a role in perceptual recognition is obvious. But the argument that familiarity is the sole or primary factor is refuted by the present results.

Another implication of the present data pertains to the research on memory for completed and incompleted tasks under ego-involved and nonego-involved conditions. The experiments in this area have generally been content with a comparison of the experimental and control groups on the mean number of completed or incompleted tasks recalled. The present findings indicate that people vary in their defenses against threat. Some favor their successes and others their failures in recall. Thus the effect of ego involvement in the tasks may not affect the group mean, but would rather affect the dispersion of individuals about this mean.

One further point deserves brief mention. It is the apparent consistency in the individual's defense against various forms of ego-threat. Different kinds of ego-threat were involved in the present study. Yet the individual who responded to one kind of threat by avoidance defenses used similar techniques against other forms of ego-threat.

SUMMARY

The present study investigated: (*a*) individual differences in response to ego-threat; and (*b*) the effect of these differences upon the perceptual recognition of ego-threatening stimuli. Two subexperiments were involved.

Experiment I was a memory for incompleted tasks. The 14 tasks were designed so that subjects would be successful on half and fail the other half. The tasks were administered to an experimental (ego-involved) group and a control (nonego-involved) group.

Experiment II determined the relation between association time and recognition threshold for various words.

The results show:

1. The main effect of ego involvement on the recall of completed vs. incompleted tasks is to increase greatly the group variance. Although the experimental group recalls relatively more completed tasks than does the control group, the most marked difference is their greater variability. The increase in variance with ego involvement reflects marked individual differences in response to threat.

2. Individuals who responded to ego-threat in Experiment I with a high success-to-failure recall ratio show considerably larger correlations between their association time and recognition thresholds for words than do individuals with low success-to-failure ratios. This result is interpreted as demonstrating the role of different ego defenses in the perception of ego-threatening stimuli.

REFERENCES

1. ALPER, THELMA G. Memory for completed and incompleted tasks as a function of personality: an analysis of group data. *J. abnorm. soc. Psychol.*, 1946, 41, 403–421.
2. BRUNER, J., & POSTMAN, L. Emotional selectivity in perception and reaction. *J. Pers.*, 1947, 16, 69–77.
3. ERIKSEN, C. W. Perceptual defense as a function of unacceptable needs. *J. abnorm. soc. Psychol.*, 1951, 46, 557–564.
4. ERIKSEN, C. W. Some implications for TAT interpretation arising from need and perception experiments. *J. Pers.*, 1951, 19, 282–288.
5. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies "Emotionality and Perceptual Defense." *Psychol. Rev.*, 1949, 56, 244–251.
6. LAZARUS, R. S., & ERIKSEN, C. W. Personality dynamics in auditory perception. *J. Pers.*, 1951, 19, 471–482.
7. POSTMAN, L., BRUNER, J., & MCGINNIES, E. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, 48, 142–154.
8. THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Teacher's College, Columbia Univ., 1944.

Received May 18, 1951.

THE INTERPERSONAL PROCESS IN PSYCHOTHERAPY: DEVELOPMENT OF A RESEARCH METHOD[1]

ALLEN T. DITTMANN
University of Michigan

In recent years most research in psychotherapy has been designed to test the general effectiveness of psychotherapy in treating patients with emotional disturbances. Criteria for effectiveness have been either ratings of general over-all outcome ("recovered," "improved," and the like) or judgments of changes in more specific perceptual and attitudinal factors ("self-concept," "positive attitudes toward others," and the like). Although these studies show that psychotherapies of widely differing orientations are usually useful to patients, they do not show what specific activities of the therapist contribute to the improvement of the patients. The research to be reported in this paper is concerned with a method for analyzing therapist behavior and its relationship to immediate progressive and regressive movement in therapy.

Two studies bearing on this problem have thus far been reported in the literature, one by Snyder (7), and the other by Keet (5).[2] These investigators have approached the problem in two different ways: Snyder by categorizing individual therapist and patient responses and comparing the consistency of the associations between the two, and Keet by "treating" experimentally induced "neuroses" by two differently oriented types of counseling interview. Both of these studies yielded important information about therapeutic interviewing, but each had its limitations as to applicability to interviews of more widely varying orientations. Snyder's design is restricted to nondirective counseling—the categories themselves were originally formulated to differentiate "directive" from "nondirective" counselor techniques. Comparable research in psychoanalytic technique would require a system of classification broader than the one Snyder used. Keet's study is the first truly experimental one to be reported in the field of psychotherapy, but it treats of general approaches, not of the specific techniques within them. The aim of the present research is to develop a method of analysis at once broad enough to handle a wider range of therapeutic approaches and specific enough to test the relative effectiveness of differing activities on the part of the therapist within each approach.

Theoretical Background

The first problem to be met in research on psychotherapeutic techniques is the selection of specific techniques which are worth studying systematically. For the present research we have drawn on the concepts of three theorists, Rogers (6), Fenichel (3), and Sullivan (9), for their ideas of what to study. These specific writers were chosen because they are well known in the field, and because they present ideas which can be tested with direct reference to interview material.

At first the contrasts among the three seem more outstanding than the congruences, especially when the psychoanalytic writers, Fenichel and Sullivan, are grouped together and compared with Rogers. As we shall see, however, the differences are mainly ones of emphasis: psychoanalytic and Rogerian concepts actually complement rather than contradict each other. In Rogerian theory the emphasis is on the person's immediate perceptions of and emotionalized attitudes toward himself and his world. In psychoanalytic theory the chief emphasis is on impulses, their striving for motor expression, and the various ways they are defended against or prevented from direct expression. Sullivan has extended psychoanalytic theory by emphasizing the interpersonal aspects of

[1] This paper is a condensation of a doctoral dissertation done at the University of California in 1950. The writer wishes to express his appreciation to Dr. Robert E. Harris for the considerable time, energy, and thought he expended in the development of this study. The writer also wishes to thank Dr. E. Lowell Kelly and Dr. Harold L. Raush for their very useful suggestions in the presentation of this paper.

[2] Since the completion of this study and the writing of this paper, another study of patient-therapist interaction has been published by Daniel V. Bergman. See Counseling, method and client responses. *J. Consult. Psychol.*, 1951, 15, 216–224.

the expression of and defense against impulses, the chief one of which is anxiety.

Rogers' emphasis, then, is on perceptions or emotionalized attitudes ("feelings"), while the psychoanalysts' emphasis is on motor expression of impulses ("behavior"). The differences in what these writers regard as effective psychotherapeutic techniques follow directly from their theoretical differences. According to Rogers the counselor should encourage free expression on the part of the client, listen closely for the immediate feelings which are expressed, and then reflect these feelings back as accurately as possible so that the client will be able to perceive himself and his attitudes more clearly. Two things are important in counselor behavior: (*a*) he must select the client's *feelings* as the material to which to respond; counselor attention to anything but feelings is by definition directive as opposed to nondirective counseling. Similarly, (*b*) the counselor must confine his activities to material which has *already been expressed;* if he tries to deal with anything which he infers to be present, but which has not yet been expressed, he is in danger of touching upon material which has been repressed, thereby stirring up resistances to counseling.

In psychoanalytic treatment such widely differing theorists as Fenichel and Sullivan agree that the therapist should help the patient expand his awareness so that he will be able to deal more directly and realistically with ideas and impulses which he has heretofore found it necessary to defend himself against. This goal is accomplished by encouraging free expression on the part of the patient of whatever comes to mind at the moment, listening for superficial derivatives of the impulse-defense struggles at the root of the difficulties, then interpreting these derivatives. Of special importance in this process is the warning (3, pp. 43 ff.) that interpretations must always deal with superficial derivatives, with material which is currently near the surface of the patient's awareness. Any attempt by the analyst to guess at and name unconscious conflicts themselves must fail, the patient being unable to accept the interpretation for the same reason that he repressed the impulses in the first place. "As is well known," Fenichel writes, "we interpret what is in the preconscious—and just a *little bit* more—which thereby becomes capable of entering consciousness" (3, p. 53). There are thus degrees of "depth" to which therapist activities might be "aimed." (For fuller discussion of this continuum as it relates to both psychoanalytic and nondirective therapy, see [2].)

The psychoanalyst, then, works with material at the surface of the patient's awareness and just a little bit beyond, while the Rogerian therapist confines his activities to material at the surface only. The material which the psychoanalyst works with comprises all of the patient's verbal associations, *plus* the patient's behavior in the therapeutic hour—his choice of subject-matter, his voice inflections, motor activity, and the like. It was Sullivan who dealt most fully with these behavioral aspects of the patient's communications, and, in accordance with his theoretical emphasis, with their interpersonal implications—with what the patient is trying to do to the therapist, what reactions he is trying to evoke by talking in this tone of voice, or by bringing up this topic at this time, or by saying these things about himself, or by asking these questions. The therapeutic situation is seen by Sullivan as an opportunity for the patient to explore the anxiety-provoking aspects of his behavior with other people, the exploring itself being carried out *with* an exemplary other person, the psychotherapist, so that the energy formerly bound up in the confusion between present and past persons can be made more available for satisfactions and security in relationships in the present.

We see, then, that all three of our writers have either explicitly or implicitly stressed the importance of three phases of activity in the psychotherapist: (*a*) the therapist's attentiveness to the communications of the patient; (*b*) the therapist's choice of feeling or interpersonal behavior as aspects of the patient's communications to which he might respond; and (*c*) the therapist's choice of the depth of the patient's awareness to which he might aim his responses.

Statement of the Problem

The general goal of this research is to develop a method for answering the follow-

ing very broad question: What specific kinds of therapist activity are associated with progress on the part of the patient toward "improvement"? The criterion of "progress toward 'improvement'" refers to an embodiment of the ideas of success in psychotherapy culled from the three theorists cited above; fuller discussion of this concept will be found in a later section along with the method of its evaluation.

It was assumed that the behavior of any psychotherapist will vary enough from time to time to permit reliable differentiation of his activities. This variation will be found not only from interview to interview, but within interviews as well. Given an appropriate criterion of therapeutic movement in the patient, the various activities of a single therapist can be compared with one another as to the amount and kind of therapeutic change.

Method

In general the procedures of the research were divided into three parts: (*a*) using categories and ratings to be described later, one set of judges evaluated an extensive sample of responses made by a single therapist in a series of successive therapeutic interviews;[3] (*b*) another set of judges evaluated the therapeutic movement shown in the responses of the patient in the same interviews,[4] and (*c*) these two groups of evaluations were compared with one another to determine the significant associations between them.

In an effort to insure independence of these evaluations, two controls were included in the design: (*a*) the judges who worked with the patient's responses were given an abridged edition of the interviews containing only the patient's responses, i.e., editions from which the therapist's responses had been deleted; (*b*) the judges who worked with the therapist's responses were given the complete interviews, but were required to evaluate each specific behavior of the therapist before reading the response which it elicited from the patient. This was accomplished by covering all portions of the typescript following the one being evaluated.

The data used

The basic data for the research consisted of the typescripts of the first 30 of a series of 43 electrically recorded psychotherapeutic interviews of the same patient. Because of the very large number of responses involved some method of sampling was necessary to reduce the data to a workable amount of material. Since the patient's presenting complaints centered around difficult relationships with specific people, it was decided to use as the sample for study only those sections of the interviews which dealt with those relationships. The sample of responses contained 965 responses by each participant and was selected independently by an assistant who also typed the abridged copies for the raters of the patient's responses.

PROCEDURES FOR CATEGORIZING AND RATING RESPONSES

Single Responses of the Therapist

Judges were instructed: (1) to rate each therapist response as to participation level, i.e., the degree of attentiveness paid by the therapist to the patient's immediately preceding communications; (2) to categorize each therapist response as follows: (*a*) response to feeling, (*b*) response to interpersonal behavior, (*c*) response to both feeling and interpersonal behavior, and (*d*) response to neither feeling nor interpersonal behavior; (3) in the case of responses categorized as (*a*), (*b*), or (*c*), under (2), i.e., responses to feeling and/or interpersonal behavior, judges were also asked to rate the therapist's response on the basis of the level of the response on a continuum ranging from superficial to deeply interpretive level. Judges were instructed to interpret depth not in terms of the therapist's intent, but in terms of the probable acceptance of the response by the patient.[5]

The objectivity of the categorizing and rating was estimated by the use of independent judges in the evaluation of 31 per cent of the therapist's responses. The scales used for evaluating therapist responses and the interjudge agreement obtained for each were as follows: judges agreed in 88 per cent of responses with respect to their assignment to the categories of response to feeling and response to interpersonal behavior. On the ratings themselves, the judges agreed exactly in 53 per cent of responses on the response to feeling scale (Pearsonian $r = .63$); 68 per cent on the response to interpersonal behavior scale ($r = .80$); and 59 per cent in the participation level scale ($r = .41$).

[3] Three judges served to evaluate therapist responses: the writer (who was also the therapist) judged all responses in the study; Dr. Timothy Leary served as independent judge of individual responses for test of inter-judge agreement of the categories and rating scales used; and Dr. George S. Welsh served as independent judge of larger units of therapist activity.

[4] Mr. Stephen S. Rauch and Mrs. Mary Darby Rauch served as judges in the evaluation of therapeutic movement in the patient's responses. The writer wishes to express special appreciation to them, not only for the ratings involved, but also for their constructive ideas in developing a workable method for evaluating patient responses.

[5] Therapist responses were also coded for ego-oriented interpersonal mechanisms as defined by Freedman, Leary, Ossorio, and Coffey (4, 1). The description of that coding system and the results of that part of the study are beyond the scope of this paper.

Scale for Rating Participation Level.

1. Therapist has not been listening to the patient or otherwise observing his activities.
2. Therapist has been listening, but his attention is directed to some relatively unimportant part of what the patient has been saying or doing.
3. Therapist is directing his response to important areas of the patient's communications, but selectively, so that he is responding to only one side of an ambivalence, for example, or trying to "make something" out of a casual remark, while neglecting the general flow of expression.
4. Therapist is "with the patient," is attending closely, and dealing with what the patient is communicating.
5. Therapist is participating "ideally," gives evidence of understanding all aspects of what the patient is communicating—all facets of ambivalences, feeling and behavior, the relation between the present (therapeutic) relationship and other interpersonal relations in which the patient is or has been involved.

Scale for Rating Level of Response to Feeling.

1. Questioning feeling. "How do you feel about it?"
2. Reflecting feeling. "You feel that you were silly for reacting that way."
3. Clarifying feeling. "It seems to have been going on less, recently, but it's hard to tell whether it's just that they haven't been crying or whether you've been getting something out of these contacts."
4. Clarifying feeling, plus just a *little bit* more. "It might be safer to make a joke out of it rather than just let the feelings come as they will."
5. So-called "deep interpretations"—reserved for those responses where the therapist is guessing at the unconscious feelings underlying the present material more from a knowledge of personality theory than from close attention to the patient's current communications.

Scale for Rating Response to Interpersonal Behavior.

1. Questioning behavior. "Why do you think you are bringing this up right now?"
2. Reflecting behavior. "All of a sudden you stop and say you have no reason to do the things you have done."
3. Clarifying behavior. "I get the impression that you're trying to get something across to me, but that you're sort of stuck."
4. Clarifying behavior, plus just a *little bit* more. "You keep bringing this up, and I wonder if it isn't intended to impress me with the seriousness of the incident, or something like that."
5. So-called "deep interpretations," referring to responses similar to those at this level on the response to feeling scale.

Larger Units of Therapist Activity

In addition to ratings of individual responses of the therapist, it appeared desirable to evaluate longer periods of therapist activity. Here arises the problem of unit size: how large should the units be, and how should they be defined? In working with the scales described above it was found that therapist activities were relatively consistent over periods of time longer than that of the individual response. Consistency of a given quality of therapist activity through a sequence of responses was therefore used as an objective criterion for separating the material into larger units. The rating scale used for evaluating these longer periods of therapist behavior was a continuum ranging from "low" to "high" quality of therapist activity. Ratings on this scale were made not from interview material itself, but rather from the ratings on the individual scales. Thus the scale for periods of therapist activity is a second order variable, being once removed from the raw data. These larger sequences of therapist activity were defined as a group of responses generally falling within the following limits:

Rating of Participation		Response to Patient's:	
Quality	Level	Feeling	Interpersonal Behavior
4	at level 5	at depth 2 to 4 *and*	at depth 1 to 4
3	at level 4 or 5	at depth 2 to 4 *and*	at depth 1 to 3
2	at level 3 or 4	at depth 1 to 5	*
1	at level 2 or 3	at depth 1 to 5	*or* at depth 1 to 5

* Responses to interpersonal behavior alone at participation level of 3 to 4 were found in preliminary work to be isolated responses so that longer sequences of therapist behavior could not be delineated on the basis of these responses.

It must be kept in mind that *consistency* of activity is the important consideration in dividing the material into these periods of therapist activity. A few isolated responses in a large period might not qualify for inclusion in that period, and yet if the therapist does not continue responding in that way long enough for another period to be identified, the context of the other responses will determine the rating for the period. It is the *context* of responses which determines both the periods themselves and the ratings of the periods on this scale.

Determining the inter-judge agreement on the scale for periods of therapist activity presented special problems. Judges here had first to separate the periods one from another, then to rate the therapist activities within each period. For the sample of the interviews selected to test inter-judge agreement on this scale one judge identified 26 periods of therapist activity, while the other found 31. The five additional periods identified by the second judge occurred within three long periods identified by the first; thus the judges agreed on the identification of 23 periods. The results are complicated, however, by those occasions where there is disagreement between the judges as to when one period ends and the next one begins. We have called this phenomenon "time lag." There were 13 instances of time lag in the sample: in seven the lag was one response; in four it was two responses; and in the remaining two it was three responses. No statistical test has been found to handle this situation, but the

agreement does not appear to be the result of chance. The agreement on the ratings themselves, moreover, substantiates this opinion: the judges agreed exactly on 19 of the 23 ratings of commonly identified sequences, disagreed by only one point on the remaining four (a Pearsonian r of .90).

Evaluation of Therapeutic Movement on the Basis of Patient Responses

The criterion of the effectiveness of therapist activities in this investigation was the direction of therapeutic movement indicated in the patient's *immediately* following responses. The word "immediately" is emphasized, because we are interested here in the *immediate progressive and regressive reactions* of the patient rather than in the more far-reaching effects of the treatment as a whole.

The definition of progress toward improvement stems from the same writers who were discussed in the theoretical background for this paper. As in their theories on technique, these writers' conceptions of the goals of treatment also differ in emphasis, but are not mutually exclusive. Rogers emphasizes increase in positive attitudes toward self or others; Fenichel stresses increase in free and realistic expressions of basic impulses; and Sullivan writes of increase in awareness of the interpersonal consequences of behavior. The goals emphasized by each are the consequence of their general theoretical orientations. Nowhere does any of them deny the importance of the goals stressed by the others.

The judges were asked to rate the immediate progressive and regressive movement in the patient's responses in terms of an embodiment of the goals of the three theorists as stressed above.[6] The rating scale used was an attempt to get at movement toward or away from these general goals. The scale had five points, as outlined below:

5—Rapid movement toward improvement
4—Gradual movement toward improvement
3—No change in either direction
2—Gradual movement away from improvement
1—Rapid movement away from improvement

The judges conceptualized this movement almost as physical movement in space: going forward or backward; going toward or away from treatment; getting closer to or moving away from the therapist; searching actively for troublesome aspects of the self or running away from the self. In simpler terms, progress versus regress, adience versus abience, integration versus isolation. The specific words which have been used here are simply illustrative of the process or direction of the movement; they were not used as the criteria for making the ratings.

[6] It is true that the judges might have allied themselves with one or another of the theories current today and thus made ratings within a concept of "improvement" more narrow than the one intended here, but the effects of such possible influence on the results of investigations like this one can only be evaluated by replication of the studies.

In order to get at the immediate movement as represented in the patient's responses, the judges were asked to rate each response serially, and to rate each response in terms of the one immediately preceding it. The ratings thus had a moving reference point. The scale itself had no absolute criterion of improvement built into it. With a scale constructed in this way, the ratings were difficult to comprehend. Plotted on a graph, however, it became possible to visualize the patient's progressive and regressive movement throughout the interview or part of an interview under consideration. Since the interview material was divided into "periods" of therapist activity, the ratings of therapeutic movement in the patient had to be summarized for these same periods. Progress responses and regress responses were accordingly counted within each period and the larger number subtracted from the smaller. For any given period an over-all rating could be given of general progress, general regress, or no change.

In the sample of patient responses selected to test inter-judge agreement on the patient progress-regress scale there was perfect agreement between the judges on 58 per cent of the responses (Pearsonian r of .69).

RESULTS

The primary goal of this study was that of determining what kinds of therapist behavior are associated with immediate progressive and regressive therapeutic movement. Table 1 shows the frequency with which the ratings of "quality" of therapist activity are associated with ratings of therapeutic movement for the same periods of the interviews. A chi-square test for the significance of the association among the events in this table is not appropriate because the entries in columns 2 and 4 do not yield high enough expected frequencies. The four columns can, however, be condensed into two columns, as in Table 2, and the chi-square then be computed. Its value is 27.7, exceeding by far the value required

TABLE 1

ASSOCIATION BETWEEN RATED "QUALITY" OF THERAPIST BEHAVIOR AND IMMEDIATE THERAPEUTIC MOVEMENT IN THE PATIENT

THERAPEUTIC MOVEMENT	NUMBER OF PERIODS IN WHICH "QUALITY" RATED OF THERAPIST ACTIVITY WAS RATED: (LOW) 1	2	3	(HIGH) 4	Total
Progressive	2	18	35	12	67
No Change	3	31	24	1	59
Regressive	7	27	8	0	42
Total	12	76	67	13	168

TABLE 2

CONDENSATION OF TABLE 1

RATED THERAPEUTIC MOVEMENT	PERIODS IN WHICH "QUALITY" OF THERAPIST ACTIVITY WAS RATED: LOW: 1 OR 2		HIGH: 3 OR 4		TOTAL	
	N	PER CENT	*N*	PER CENT	*N*	PER CENT
Progressive	20	22	47	59	67	40
No Change	34	39	25	31	59	35
Regressive	34	39	8	10	42	25
Total	88	100	80	100	168	100

for the .001 level of confidence with two degrees of freedom.

From these results it may be concluded that there is a significant association between high "quality" of therapist behavior and progressive movement, i.e., between high level of participation *and* response to both feeling and interpersonal behavior, on the one hand, and progress on the other. By implication it may also be concluded that there is a significant association between low "quality" therapist behavior and regressive therapeutic movement, i.e., between low level of participation *and* response to either feeling or behavior, on the one hand, and regress on the other. Closer inspection of Table 2, however, reveals that this latter association has not been shown so clearly: whereas 59 per cent of periods rated high in "quality" are associated with progress, only 39 per cent of periods rated low in "quality" are associated with regress.

The association thus far shown is between "quality" of short sequences of therapist behavior and therapeutic movement in the patient. Therapeutic movement as defined in this study is a simple variable, the ratings having been made on the raw data of the interviews themselves. "Quality" of therapist behavior, on the other hand, is a complex variable, since the ratings were based on: (*a*) the therapist's participation level, i.e., the degree of the therapist's attentiveness to the patient's communications, and (*b*) the therapist's choice of responding to the patient's feeling or his interpersonal behavior. Both of these considerations are reflected in the ratings on "quality" by the very definition of the rating scale: high "quality" ratings were to be given to periods when the therapist is usually participating at a high level and responding to *both* feeling *and* behavior, while low "quality" ratings were to be assigned to periods when the therapist is generally participating at a low level and responding to *either* feeling *or* behavior. These instructions resulted in a high relationship between participation level and response to feeling and/or behavior, but the relationship was not unity. Further analysis was therefore undertaken to study the association between response to feeling and/or behavior and therapeutic movement, with variations in participation level held constant. Participation level rather than feeling and/or behavior was chosen as the variable to be controlled because of the following theoretical consideration: all theorists agree that good psychotherapy requires a high level of participation by the therapist during the interviews, but they disagree on the particular aspects of the patient's communications to which therapists should respond. Rogers holds that therapists should respond to *feeling only*, while the psychoanalysts, especially Sullivan, say that therapists should respond to interpersonal behavior as well as feeling.

In the following analysis the therapist's participation level was held constant by studying only those therapist responses rated high on this variable. Thus instead of grouping responses into "periods in which therapist activity was rated high in 'quality,'" as in Table 2, we now grouped responses into "periods in which the therapist was generally responding to *both* feeling *and* behavior." Similarly the classification of responses into "periods in which therapist activity was rated low in 'quality'" was now changed to "periods in which the therapist was generally responding to *either* feeling *or* behavior." Individual responses of the therapist could now be examined taking full account of the general behavior of the therapist during the periods in which they occurred.

Therapist Response to Feeling Only

Table 3 presents the frequency for high participation level responses to *feeling only*, which occurred (*a*) during periods in which the therapist was responding to *either* feeling *or* behavior, and (*b*) during periods when he was responding to *both* feeling *and* behavior. The periods in which these responses occurred are tabulated, as in Table

2, according to the therapeutic movement of the patient. Note that when the therapist is responding generally to either feeling or behavior, the majority of responses to feeling (59 per cent) are associated with "No Change" in therapeutic movement on the part of the patient; the next largest proportion of these responses (26 per cent) is associated with progress. By contrast, when the therapist was generally responding to both feeling and behavior the majority of responses to feeling (67 per cent) were associated with progressive movement. In other words, response to feeling only was more "efficient" with respect to therapeutic movement when it occurred in the context of response to feeling and interpersonal behavior.

TABLE 3

ASSOCIATION BETWEEN HIGH PARTICIPATION LEVEL THERAPIST RESPONSE TO FEELING ONLY AND THERAPEUTIC MOVEMENT IN THE PATIENT

RATED THERAPEUTIC MOVEMENT	FREQUENCY OF RESPONSES TO FEELING WITHIN PERIODS WHEN THERAPIST WAS GENERALLY RESPONDING TO: *Either* Feeling *or* Behavior N	PER CENT	*Both* Feeling *and* Behavior N	PER CENT	TOTAL N	PER CENT
Progressive	23	(26)	33	(67)	56	(41)
No Change	51	(59)	15	(31)	66	(49)
Regressive	13	(15)	1	(2)	14	(10)
Total	87	(100)	49	(100)	136	(100)

Therapist Response to Interpersonal Behavior Only

In the case of individual responses to interpersonal behavior only, a different set of associations obtains. Table 4 presents the appropriate data, again using only those responses rated high in participation level. When the therapist was responding generally to either feeling or behavior, the majority of responses (67 per cent) to behavior only were associated with "No Change" in therapeutic movement. This finding is similar to that for responses to feeling only (cf. Table 3). Responses to behavior only differ from responses to feeling only, in that in this context, if they are associated with any change, the change is regressive. When the therapist was responding consistently to both feeling and behavior, the largest proportion of responses (57 per cent) was again associated with progressive movement. Summarizing the findings from Tables 3 and 4, it would appear that while it was more "efficient" for the therapist to center his activities on feeling only than on behavior only, it was still more efficient for him to avoid the isolation between the two, i.e., to direct his therapeutic activities toward both of these aspects of the patient's communications at once.

TABLE 4

ASSOCIATION BETWEEN RESPONSE TO INTERPERSONAL BEHAVIOR ONLY AND THERAPEUTIC MOVEMENT IN THE PATIENT

RATED THERAPEUTIC MOVEMENT	FREQUENCY OF RESPONSES TO BEHAVIOR WITHIN PERIODS WHEN THERAPIST WAS GENERALLY RESPONDING TO: *Either* Feeling *or* Behavior N	PER CENT	*Both* Feeling *and* Behavior N	PER CENT	TOTAL N	PER CENT
Progressive	4	(7)	34	(57)	38	(33)
No Change	37	(67)	17	(28)	54	(47)
Regressive	14	(26)	9	(15)	23	(20)
Total	55	(100)	60	(100)	115	(100)

Therapist Responses to Both Feeling and Interpersonal Behavior

The conclusions drawn from the findings in Tables 3 and 4 are further strengthened by consideration of individual responses to both feeling and behavior. The figures for these responses are presented in Table 5, still including only high participation level responses. As would be expected, not many of these responses fall in periods when the therapist is responding to either feeling or behavior. They do, however, make up the majority of responses during periods when

TABLE 5

ASSOCIATION BETWEEN RESPONSE TO BOTH FEELING AND BEHAVIOR AND THERAPEUTIC MOVEMENT IN THE PATIENT

RATED THERAPEUTIC MOVEMENT	FREQUENCY OF RESPONSES TO BOTH FEELING AND BEHAVIOR WITHIN PERIODS WHEN THERAPIST WAS GENERALLY RESPONDING TO: *Either* Feeling *or* Behavior N	PER CENT	*Both* Feeling *and* Behavior N	PER CENT	TOTAL N	PER CENT
Progressive	3	(14)	65	(52)	68	(46)
No Change	16	(77)	43	(34)	59	(40)
Regressive	2	(9)	18	(14)	20	(14)
Total	21	(100)	126	(100)	147	(100)

the therapist is responding consistently to both feeling and behavior, and their pattern of associations with therapeutic movement is the same as those for responses to feeling only and responses to behavior only: most responses to both feeling and behavior (52 per cent) fall in periods of progressive movement.

The Depth of Therapist Responses

Thus far in this analysis of individual responses, results have been presented solely in connection with the *categories* of response to feeling and response to interpersonal behavior, without regard for the level or depth of the therapist's responses. An analysis of the ratings of depth is now presented in order to throw light on the theoretical issue of the relative advisability of superficial versus deep responses by the therapist to the material presented by the patient. Rogers holds that the therapist should confine his activity to material of which the patient is fully aware, while the psychoanalysts, particularly Fenichel, take the position that the therapist should "interpret what is in the preconscious —and just a *little bit* more."

Table 6 presents the frequencies of responses categorized as to depth into two categories as they are associated with periods

TABLE 6

ASSOCIATION BETWEEN DEPTH OF THERAPIST RESPONSE AND THERAPEUTIC MOVEMENT IN THE PATIENT

	NUMBER OF THERAPIST RESPONSES					
	SUPERFICIAL		DEEPER		TOTAL	
	N	PER CENT	*N*	PER CENT	*N*	PER CENT
Periods of Progressive Movement	78	(37)	84	(44)	162	(41)
Periods of No Change	92	(44)	87	(46)	179	(45)
Periods of Regressive Movement	38	(19)	19	(10)	57	(14)
Total	208	(100)	190	(100)	398	(100)

of progressive and regressive therapeutic movement. Again the analysis was confined to responses of high participation level. The category called "superficial" in this table includes responses rated "questioning" and "reflecting"; the category "deeper" includes responses rated from "clarifying" to "deep interpretation." It should be noted parenthetically that only about 2 per cent of responses in the data used for this study were rated "deep interpretation," so that the second category in Table 5 actually includes responses rated "clarifying" and "clarifying plus just a *little bit* more." The chi-square test for the significance of the differences among events in this table is 5.45, which approaches significance at the 5 per cent level of confidence with two degrees of freedom. It is to be noted that while periods of no change account for almost half of all responses both "superficial" and "deeper," regressive periods contain almost twice the proportion of "superficial" responses as compared with the proportion of "deeper" responses. Although the significance level of the difference between these proportions is not high, the direction of the difference is more in accord with the recommendations of Fenichel than with those of Rogers.

DISCUSSION

An obvious fact serves to mitigate the degree to which the results of this investigation can be generalized, namely that the data derive from only one patient and one therapist, even though the sample of their behavior is a large one. While the results which have been presented are indeed provocative for theory testing, it is certainly questionable whether they should serve as a guide for psychotherapists.

There is, however, another side to this word of caution: while it is true that the results of this particular study are highly dependent on the vicissitudes of the particular patient and therapist under study, it is also true that the patterns of interaction between them occur again and again between other therapist-patient pairs—between this therapist and other patients and between other therapists and other patients. We cannot assume that this therapist and this patient are completely representative of therapists and patients in general, but neither can we assume that any one therapist and one patient are completely different from all others—for if we did, we could not generalize no matter how many data we collected. The most fruitful way of continuing this research is to expand the sample of the various patterns of interaction identified in this study by collecting large samples of the behavior of other therapist-

patient pairs, then to study the conditions under which each interaction pattern can be generalized to other pairs.

SUMMARY

1. With a view to clarifying the differential effectiveness of "ideal" therapist behavior as defined by three current theories (nondirective and two versions of psychoanalytic), the following general question was asked: What specific kinds of therapist activity are associated with progress on the part of the patient toward "improvement"? The data were the typescripts of the first thirty of a consecutive series of recorded psychotherapeutic interviews.

2. Therapist responses were rated as to participation level, i.e., the degree of attentiveness paid by the therapist to the patient's communications.

3. Therapist responses were also categorized as follows: (*a*) response to feeling; (*b*) response to interpersonal behavior; (*c*) response to both; and (*d*) response to neither. Responses categorized into the first three of these categories were then rated as to depth on a continuum from superficial to deeply interpretive. Therapist responses were also grouped into larger units on the basis of consistency of the activities categorized and rated in the manner described above.

4. Independent ratings were made of the progressive and regressive movement reflected in the responses of the patient.

5. Progressive therapeutic movement was found to be associated with (*a*) high level of participation on the part of the therapist; *and* (*b*) response to either feeling or interpersonal behavior or both, *providing* the response occurs in the context of consistent response to *both* feeling *and* behavior.

6. Progressive therapeutic movement was also found to be associated with therapist responses which were slightly deeper than pure "reflection."

REFERENCES

1. COFFEY, H. S., FREEDMAN, M. B., LEARY, T. F., & OSSORIO, A. G. Community service and social research—group psychotherapy in a church program. *J. soc. Issues*, 1950, *6*, 1–65.
2. COLLIER, R. M. A basis for integration rather than fragmentation in psychotherapy. *J. consult. Psychol.*, 1950, **14**, 199–205.
3. FENICHEL, O. *Problems of psychoanalytic technique*. New York: Psychoanalytic Quarterly, 1941.
4. FREEDMAN, M. B., LEARY, T. F., OSSORIO, A. G., & COFFEY, H. S. The interpersonal dimensions of personality. *J. Pers.*, in press.
5. KEET, C. D. Two verbal techniques in a miniature counseling situation. *Psychol. Monogr.*, 1948, **62** (7), Whole No. 294.
6. ROGERS, C. R. *Counseling and psychotherapy*. Boston: Houghton Mifflin, 1942.
7. SNYDER, W. U. An investigation of the nature of non-directive therapy. *J. gen. Psychol.*, 1945, **33**, 193–223.
8. SNYDER, W. U. The present status of psychotherapeutic counseling. *Psychol. Bull.*, 1947, **44**, 297–386.
9. SULLIVAN, H. S. Conceptions of modern psychiatry. *Psychiatry*, 1940, **3**, 1–117.

Received May 28, 1951.

CASE REPORT

THE CULTURAL BACKGROUND OF THE PATIENT AS PART OF THE PHYSICIAN'S ARMAMENTARIUM

Four Case Studies

HARRY AUSTRYN SAVITZ, M.D.
Tufts Medical School

IN the works of Plato in the *Dialogue Charmides* we find the following amazing passage:

"But our king Zamolxis," said he, "being a god, says that, as it is not proper to attempt to cure the eyes without the head, nor the head without the body, so neither is it proper to cure the body without the soul; and that this was the reason why many diseases escape the Greek physicians, because they are ignorant of the whole, to which attention ought to be paid; for when this is not in a good state, it is impossible for a part to be well."

Yet it is only recently that the medical profession has finally accepted the interrelationship of mind and body and has developed the so-called psychosomatic medicine. It seems to me that many ailments still escape us at the present time as they did in the time of the Greek physicians, for our knowledge of somatic disturbances is far more advanced than our understanding of the psyche of the patient.

The modern doctor's bag contains a number of drugs and ampules to meet practically every drastic emergency. His armamentarium also contains some tools of investigation to aid in diagnostic purposes. Tissues and fluids are subjected to the finest chemical analysis possible. For it is only in such a manner that we begin to unravel the secrets of nature that are concealed in the living organism. Now if the cultural background of his patient is also stored in the physician's mind it will often help him in understanding and analyzing his patient's troubles, and thus arrive at a right diagnosis. At times he may utilize this cultural background as a rapid, rational therapeutic measure. There are cases, as we shall show, where the rational processes of the patient are blocked either by fear or some other emotional complex. Here is a method by which an approach is made to the reasoning processes of the patient by way of a detour—through the open door of the underground cultural storehouse. The ideal physician, then, will not only sympathize with his patient, but as in art, he will possess empathy for him. Whereas sympathy is a feeling with a subject—what the Germans call "mitfühlen"; he will also have empathy or feeling into "einfühlen." By understanding thoroughly the cultural background of a patient the physician feels into the very depths of the mind of a patient. He understands his intricate complexities and so can help his patient to untangle himself from his web of troublesome perplexities that the patient spins around himself.

The following are a group of four cases from my private practice which illustrate the value of understanding the cultural background of a patient.

CASE I

Mrs. R.C., age 60, housewife, came to see me on April 9, 1947. Her chief complaint was nervousness of one year's duration. Her present illness had begun about one year before when she would awaken suddenly from her sleep and find her body muscles stiff, particularly in the back of the neck. There were no recent loss of weight and no other symptoms. She denied having any mental worries.

P.H. Was born in Russia. Since the age of seven she has been residing in the United States. Thirty-five years ago she had pulmonary tuberculosis for which she was confined to a sanatorium for ten months. No previous operations. Six years ago she fell and sprained her right hip.

M.H. Married thirty-five years. Has one son living and well. Her husband is also living and well.

F.H. Father died of heart trouble at the age of forty-five. Mother died of pulmonary tuberculosis at the age of thirty-one. One sister is living and well.

Head. Claims that she feels dizzy, no headaches.

C.R. Has no cough. No precordial pain. Claims that she sweats somewhat at night.

G.I. Appetite good. Bowels are regular with the aid of laxatives. No G.U. symptoms. Says weight was over 200 pounds thirty years ago.

Physical Examination. Revealed a well-developed, well-nourished woman; weight—164 pounds; throat—clear; eyes—negative; eye-grounds—negative; ears—negative; heart was not enlarged, rate—72, rhythm—regular, no murmurs; BP 160/80; lungs—slight dullness over right apex; no rales; abdomen—negative; reflexes—normal. In other words, the physical examination was essentially negative.

In spite of the fact that she denied having any worries, a detailed history of her life revealed the following story: Thirty-five years ago she had pulmonary tuberculosis for which she was confined to a sanatorium. On leaving the sanatorium she was told not to have children as it might reactivate her tuberculous lesion. In spite of this medical advice she took a chance and became pregnant and gave birth to a son. She gave her son the best education possible in spite of financial stress. He graduated from law school and passed the bar. During the second World War he volunteered and served in the United States Army. Upon his discharge from the army he was married and made his home in the West.

Upon questioning as to whether she sees him, she replied, "He lives too far." In telling about her son the patient spoke reluctantly and without feeling. On further pressing of the question of how often she visits him, she finally said, "They don't want to see me!—I lost him." When asked to explain she said that he had married out of his faith and was lost. The patient then became silent. Here I felt that the first indications of the cause of her symptoms were brought to the surface. Here was a mother who risked her life in order to have a child. She gave all to him—in care and education. Finally, in intermarrying—in view of her orthodox beliefs—she considered him lost.

An attempt was now made to integrate that problem in her emotional thinking and make her feel that she still possessed him. At first an attempt was made to appeal to her maternal instinct. I asked her if she would feel better if she had received a communication from the War Department telling her that her son was missing in action. At first there was no response, but on further questioning she remarked, "If you insist, I don't know." In other words, this episode, to her, was equivalent to death; and what was worse, death with disgrace instead of glory. A final appeal was made by telling her the following story. The patient was told to listen carefully to every word for she would be questioned about it afterwards. (The story "One Must Ask for Mercy," is from the Hebrew by a Palestinian author, the late Ben Eliezer.)

There was a small town in Russia which was famous for its Rabbi. He was a scholar and an authority in the law as well as a saint in the mode of living. Many large cities were competing to obtain him as their spiritual leader, but to no avail. His oldest daughter, who was married to a merchant, lived in the capital of czarist Russia, St. Petersburg. Her two children, a son and a daughter, attended the Imperial University. While there, they were involved in a revolutionary plot. A plot that was planned against the life of the Czar himself. They were caught and it was well known in those days that the sentence would be death. The parents spent money and effort but to no avail. All doors were closed to them. When all attempts had failed, the daughter came rushing home to seek aid from her father. On seeing her, the old man sensed that his daughter was in trouble and said to himself, "When all is well with children you neither see them nor seldom hear from them; but when they are in trouble they hurry home." (Here my patient became very attentive and completely absorbed in the story.) With tears in her eyes the daughter told him of her plight.

"But what can I, a Rabbi, do in such a case?" the father asked.

Then the daughter proceeded to tell him that in the capital there was a man who was very influential among the dignitaries of the Royal Court, and that only he could save their lives.

"This man is very stern and unapproachable, but perhaps you—an old Rabbi—can prevail upon him for help."

The poor, old Rabbi made the long journey to St. Petersburg and finally located and gained an interview with this influential person. The Rabbi

told him his tale of woe—of his two grandchildren who were facing death.

"What was the crime?" asked the man.

"Revolution," answered the Rabbi.

"For every crime one can find an excuse," shouted the man, "but revolution, why that undermines the very foundation of government!"

The Rabbi lowered his head in sorrow, not knowing what to say.

The man continued, "The Czar, his Excellency, can find no excuse for a rebel nor does the fact that one comes from a great family help the matter. Very often members of the royal family were condemned and executed."

The Rabbi listened, his face white as snow and his eyes seemed to express the agony of hopelessness.

Here the man tapped him on the shoulder and said, "Even so, one must ask for mercy. Meet me tomorrow at the Royal Court and I shall attempt to appeal the case for you."

The man fulfilled his promise. He made a plea and as a result the children were set free.

Having fulfilled his mission the Rabbi was about to return to his home; but, he felt that he must first go to the man and express his gratitude for saving two lives. He entered the man's home and when the man saw him, he greeted the Rabbi cordially and asked him to follow him. The Rabbi followed him from one room to another; finally he took out a key, opened a door, and again said, "Follow me." He locked the door behind them. This was his private library. To the amazement of the Rabbi, he found the shelves filled with tomes of ancient Hebrew literature—The Sacred Writings, The Mishnah, The Talmud.

"Rabbi," said the man, "This is my private sancta-sanctorum. No stranger has ever crossed this threshold. It is here that I find relaxation from my daily routine. Rabbi, I was born a Jew; but I am converted and very few know about it. Do you think, Rabbi, that there is any chance for me?"

The Rabbi looked at him and said, "For every sin one may repent on the Day of Atonement and the sin may be absolved; but, conversion, why that destroys the very foundation of religion!"

The man turned pale.

Then the Rabbi tapped him on his shoulder and said, "Even so, one must ask for mercy. . . ."

At the end of the story, the woman burst into tears and laughter, at the same time. She got up and said, "I am very happy that I came to see you." Nothing was said after that.

She was given a simple tonic. Two weeks later she called me on the telephone and said, "I have been sleeping nights and feel much better. Do you think I should refill that medicine?" She was told that she no longer needed it.

The patient visited me again on July 11, 1947. At that time she was well-adjusted to her problem. On questioning her about her son she said that he was very happily married and living a life of his own.

"My husband is visiting him now and when he returns I shall visit him myself."

The reason for this office visit was an attack of paroxysmal auricular fibrillation which cleared up in a few days.

This case was discussed with a prominent Boston psychiatrist. On completion of my story he asked in amazement, "How did you find a story that could fit the case so perfectly?"

"I didn't," I replied, "I found a patient whose case fitted a story which I had read."

Case II

Mr. F.P.N., age 51, came to my office on November 3, 1936. His chief complaint was pain in the chest, coming on with exertion and relieved by rest. Pain radiated to the left arm.

P.H. Irrelevant except for pneumonia at the age of 30.

F.H. Father died at the age of 51 of pneumonia. Mother died at the age of 47 of cancer.

M.H. Married, has three children living and well.

Physical Examination. Revealed an obese, keen, emotionally tense, business man of about 51 years of age; weight—193 pounds; vital capacity 3,000 cc; hemoglobin 95%; blood pressure 200/102; pupils were equal and reacted normally; throat—clear; there was a slight pulsus alternans; lungs—no rales; dull or bronchial breathing; abdomen—no masses or tenderness, liver and spleen were not felt; heart was not enlarged, action was regular and rapid, sounded of good quality, no murmurs were heard; EKG tracings were not remarkable except for left ventricular preponderance.

Diagnosis of hypertension, hypertensive heart disease, and angina pectoris was made.

The patient was advised to go to bed for three weeks. He was given Aminophyllin 1 tablet TID and nitroglycerin PRN under the tongue. He was also advised to lose some weight.

At the end of three weeks he was allowed to get up but was advised to retire from his business. He continued to have some precordial pain even on rest for he was tense and aggravated. Whatever one told him to do his reply was, "What difference does it make? In six months I shall be in Wakefield." (There was a family lot in the cemetery in Wakefield.) With this idea in mind he resisted any words of encouragement and would continually interrupt with the above remark. It was felt that if we could overcome his resignation and instill some hopefulness in him it would help his condition.

This man was a very devoted father and husband. His wife had an innate religious instinct. With this as a start, the following attempt was made one day.

Instead of discussing his condition with him, I told him about the romance of the rebuilding of Palestine. How a new life was starting there; and how Hebrew was rejuvenated; and how Jews were continuing to create a literature which is a continuation of the Old Bible. As he became absorbed in the pioneering drama, I interrupted the story by saying to him, "Do you know what I would like you to do? It would also please your wife very much, for you know her deep religious emotions. I would like to take a trip with you in five years from today, to Palestine. As your guide I would point out to you the numerous historical places mentioned in the Sacred Writings. You would see the romance of pioneer life in action. You would gain a great many things which you missed during your busy life. You would also gain an understanding in the character and make-up of the Jewish people. What do you say to that?"

Absorbed in this dissertation, he said, "I agree."

"Remember," I said in all seriousness, "I really mean that."

His reply was, "Let's shake on that, and my wife will certainly be pleased to hear about it."

"Remember the date." I said, "December, 1941!"

"I shall make all the arrangements." he replied.

This suggestion took root and it worked well. I called it "a five-year plan for angina pectoris."

At the end of six months, on one of my routine visits, he said, "Four and one-half years more, Doctor."

A year later he said, "Three and one-half years more to go." And so he looked forward to that goal. In December, 1941, I said, smilingly, "What about the trip?"

And he replied, "How can we go now, with hell breaking loose all over the world?"

So we agreed to let the plan go for another five years—until the world again would return to normal.

Meanwhile, another tragedy occurred. In 1942, his youngest son, who resided in the West, accidentally shot himself while cleaning his rifle. His married daughter received the sad news on the telephone at four A.M. She, in turn, called me at that hour to go and break the news to her father who was suffering from as severe an ailment as angina pectoris. As a physician, I felt it was my duty—one which I could not shirk. At that time, the patient was retired and lived in an apartment hotel. I knew that his usual habit was to rise early in the morning and sit in the lobby. I told the daughter to meet me there, and we would go together to break the sad news. I entered the lobby followed by the daughter and as he saw me enter he said, "What happened?"

To which I said, "I have some sad news to tell you, but first I want to ask you for a favor. Cooperate with me so as to save your wife, you know that she is emotionally very tense."

"What happened?" he asked again.

I answered, "Your son met with an accident."

"How? When?"

I said, "I really don't know the details, he may be dead. But what I am concerned with is your wife—how she will take it. Please cooperate with me and help me."

He began to turn pale and he was given nitroglycerin and left with his daughter while I went upstairs.

I said to the wife, "I have some sad news for you, your son met with an accident. But what I am concerned with is your hus-

band. You know that he has a serious heart condition and the slightest exertion or emotion is liable to prove fatal. I want you to cooperate with me to save him."

"But what happened?"

Again I repeated, "Your son met with an accident. He may be dead. I don't know the details. But what I am concerned with is your husband. I want to save him and I want to solicit your aid in cooperating with me." As I talked with her the husband came upstairs with his daughter.

With tears in his eyes he asked her, "How do you feel?"

"Never mind me," she replied, "how do you feel?"

Here, again, is a case which utilized the background of the patient; two five-year plans were drawn up to relieve the fear of eminent death by angina. Knowing the background and the devotion of the patients for each other, it was utilized in another emergency—in a sad tragedy which could have been made even worse.

Case III

The following case illustrates how utilizing the cultural background of the patient helped to overcome a hospital phobia in a patient who needed an emergency operation.

On the evening of March 21, 1932, I was called to see an elderly sick man on consultation—a man whose diagnosis was problematic.

Patient S.F., age 76, whose chief complaint was abdominal pain of three days' duration. This pain was generalized throughout the abdomen and would last from fifteen to twenty minutes and then vanish. There was slight nausea at first, but no vomiting. Two days previously the patient felt worse because of abdominal distress and frequent attacks of generalized abdominal pain. Patient was constipated so he took "bitter salts" with no relief. In fact, it made him feel worse. The next day he got out of bed for a few hours. His condition was not improved. The generalized abdominal pain disappeared and became more localized in the right lower quadrant. One day previously he had taken a glass of milk and then had vomited considerably. Patient was much relieved after vomiting. A physician was called in the day previously. He gave the patient some pills, with only slight relief. I was told that the physician had said it was urethral colic. The next day the patient vomited again and there was considerabe pain in the right lower quadrant; as a result I was called in to see him.

P.H. Born in Poland in 1856. Has resided in the United States for only eleven years. The rest of the history is irrelevant as he was always well.

F.H. Father died at the age of 85. Mother died during childbirth when the patient was quite young.

Physical Examination. Revealed an old man of the patriarchal type, with a long gray beard and skull cap on his head. His face was flushed; temperature was 100.5; pulse —100; respiration—24; eyes showed marked arcus senilis, both lenses showed opacities; lungs—there were rales at both bases, breath sounds were emphysematous, there was a soft blowing systolic murmur; abdomen—showed tenderness over McBurney's point; skin over the right lower abdomen was more sensitive than the left and there was pain on rebound. I felt that it was a case of acute appendicitis and urged an immediate operation.

As soon as the patient heard the word "operation" mentioned he turned his head towards the wall and refused even to discuss the matter. He claimed that it was his wish to die in his own home, in his own bed. At that time his two little rooms were filled with his sons and daughters, his grandchildren and other relatives, each of them pleading with him to listen to the doctor's advice, but he refused.

As a last resort, I entered the room of the patient and speaking to the other members of the family, I said, "Kibud Av." (Those are the first two words of the commandment, "Honor thy Father.") "Honor thy father. Please don't argue with him. You must respect his wishes. If it is your father's will not to go to a hospital, you must respect his wish. Please leave the room and let me talk to him." Then I turned to the old man and said, "Grandfather, the hour is late and I must go; but, first, I wish to ask you to do me a favor. Tomorrow when they lower

you into your grave and, as according to Jewish lore, the Angel will ask you your name and then perchance if He should ask you to mention the names of those who contributed to the cause of your death, would you please do me the favor of telling the Angel that it was not the fault of the physician. That according to the best of his ability he diagnosed the case and advised you to go to the hospital, but that you yourself refused to go. Goodbye and good luck."

"Just a moment," replied the patient.

But I said that I was sorry—that it was late and that I must leave.

"Just one second," he pleaded, "I want to ask you—does a person die from this illness?"

"No," I replied, "if you carry out the advice of the physicians and surgeons. Otherwise it may prove fatal."

He then said to his wife, "Give me my 'Arba Kanfoth'" (fringed, religious garment). And he went to the hospital.

He was admitted at 8:00 P.M. that evening and operated on the same night. A gangrenous appendix was removed. He made an uneventful recovery.

Here is a case where a patient born in Poland still retained the picture of the hospital in his native town—a dreadful place where patients were sent as a last resort, in terminal stages of their illness, and from which very few returned. By appealing to his deep religious background, which contained both law and lore, it was possible to break his resistance to his hospital phobia and proceed with modern medical treatment. As a matter of fact, one year later, on August 25, 1933, under localized anesthesia, a cataract was removed from his right eye and on December 9, 1935, a nucleation of the left cataract was performed.

Hospital records, no matter how complete in detail, lack the human element—as in this case, the fear of going to a hospital. However, without this human element there would have been no hospital record.

CASE IV

The following case illustrates an even more perplexed complex—that of a broken-hearted mother who fixed the guilt of the death of her daughter on a physician and the only way that she could be made to reason was by an appeal to a traditional Jewish ritual.

On one of New England's stormy winter days in February of 1933 I was called to see a man with an acute pharyngitis. Treatment was outlined to him and the man then said, "Will you please examine my wife as she is suffering from a heart attack."

Physical Examination. Revealed a woman, aged 56, with a moderate tachycardia; rate —120, heart was not enlarged, sounds were regular, no murmurs were heard; lungs—no rales; blood pressure 140/70; thyroid was not palpable, no exophthalmus present. Patient appeared tense and anxious. There were no signs or symptoms of hyperthyroidism. I prescribed triple-bromides and told her to take a dram TID. I reassured her that she had no heart trouble but that she was suffering from some emotional upset and that was the cause of her rapid heart action.

"Upset!" she exclaimed, "Why Doctor, my daughter was just killed by a doctor!" Here she mentioned the name of a reputable Boston surgeon.

"What do you mean by that?" I asked, and she told me the following story:

Her daughter, age 26, had been stricken ten days previously with signs and symptoms of acute appendicitis. A surgeon was called in. He diagnosed the case correctly and took her to the hospital. She was operated on and an acute inflamed appendix was removed. At the end of eight days she was brought home. A few hours later, while walking into her bedroom and looking over some gifts that had been sent to her, she dropped dead.

"Why, madam," I said, "the doctor did not kill your daughter. He did everything that was right and your daughter died of one of the complications that sometimes follow an operation."

She interrupted, screaming, "You doctors," and jumped from her chair in a fit of hysteria. "You doctors have a union. You never talk against each other but you know very well in your heart that he killed her! If I had had Dr. ..... this thing would never have happened."

With this she collapsed in her chair, weeping.

I made several attempts to calm her, telling her that a similar episode had happened to the son of a state official whose death notice had recently appeared on the front page of all the Boston papers, but to no avail.

Again she said, "You doctors have a union, but you know the truth. That he killed her! Why didn't he kill me? I have lived my life! She was my source of joy and happiness!" Here she narrated, with tears in her eyes, the story of the life of her daughter, how she excelled in her studies, how beautiful she was, and with this she rushed out and brought in a picture of her. Then she described the funeral, etc.

Here is a mother whose heart was broken and bleeding over the premature death of her only daughter and she had fixed the guilt of death on the surgeon who had done his duty and performed the operation. Any attempt to comfort her met with further association of me as a member of "the union" and thus a partner to the crime. I tried to counteract this by appealing to the cultural, religious background of the patient. I interrupted the conversation suddenly by saying to the woman, "Pardon me, but I forgot to ask your husband a very important question." I then said to the husband, "Do you wash your hands after going to the bathroom, according to the Jewish ritual law?"

"Yes," he replied.

Meanwhile, the woman remained with her mouth open, as if in the middle of a sentence, while tears were still flowing from her eyes.

I then said to the husband, "Do you recite the prayer, 'Asher Yozar,' after you wash your hands?"

"To tell the truth," he said, "sometimes I do and sometimes I don't."

"Now," I said, "when you do recite it, do you know what it means?"

And he answered, "I must confess that I don't."

"Well," I said, "I shall recite the prayer for you, as it is one of the finest medical descriptions to be found anywhere."

"What does it mean?" he asked.

I started the recitation: "Blessed art Thou, O Lord our God, King of the universe, who has fashioned man in wisdom, and created in him many orifices and tubes. It is revealed and known before Thy glorious throne, that if but one of these be opened or stopped, it would be impossible to exist and to stand before Thy presence. Blessed art Thou, O Lord, who healest all flesh and doest wondrously." As they both listened attentively I finished by saying, "One of those tubes closed—that is the cause of your daughter's death."

The husband turned to his wife and said, "Do you hear? The doctor is right. So it is written and it is the fault of no one."

Here I felt that my mission was performed and I had helped to sooth a mother's broken heart. I put on my coat and was about to leave when the woman turned to me and said, "Please, doctor, cool off, you are sweating. It is cold and stormy outside. You may catch cold. We need you. You are a friend in need."

Discussion and Summary

By knowing well the cultural background of the patient, the physician can view the universe through the eyes of the patient. He will then be able to understand the complexities of the texture and pattern of the patient's reasoning. A patient's actions and motivations become clear and understandable. A lack of understanding of the deep cultural background will often mislead a physician and cause him to draw a wrong conclusion, as illustrated by the following episode:

An elderly patient of mine, suffering from advanced Buerger's disease, was admitted to the hospital for amputation of the involved leg. I witnessed the operation which was done under spinal anesthesia. In the tense silent atmosphere of the operating room, as the surgeon was sawing the bone, the patient was heard mumbling in a chanting voice as if accompanying the rhythm of the saw. Whispering to the young resident anesthetist I inquired about the patient's condition. He replied, "His pulse and blood pressure are good; but it seems as if the spinal anesthetic has affected his mind." I turned my ear to the patient and I heard him humming in Hebrew the words of an ancient

prayer, "God, who performs miracles, who is the creator of medicine, etc. . . ." Does this indicate that the patient's mind was affected, or does it not point out that the young anesthetist was misled because of lack of knowledge of the patient's deep cultural background.

If we wish really to understand a patient's mode of thinking we must delve deeply into his history, culture, and tradition. In mental ailments, as in organic disease, we must understand the natural processes and work with them, not against them. This is true of all peoples and all nationalities. This is particularly true in the case of the Jew, who has a long history, a continuous culture, a vast literature, and an endless chain of traditions. All these factors modify the pattern of one's thinking. The reasoning processes are like long equations with several unknowns, where each symbol has to be worked out in order to solve the equations. Nor is there a slide rule formula in medicine which would apply to all patients. Each one has to be studied individually and an estimate made of the entire personality.

Knowing the cultural background of the patient we can utilize it also as a therapeutic measure, for it permits the physician to penetrate into the very depths of the subconscious mind of the patient and bring it to surface. The cultural background acts as a lever and will penetrate deep into the subconscious mind. The more the physician knows about the cultural background of the patient the longer the arm of the lever is for lifting and bringing his point into the focus of the patient.

These four cases are not merely simple empirical incidents wherein the physician is called upon to give emotional support. The preservation and reproduction of the human voice or the recording electrocardiographically of electrical potentials of the heart muscle appear very simple superficially when once seen performed; but an analysis of these mechanical devices will reveal intricate basic scientific principles. So an analysis of these cases will reveal deep-rooted psychological problems and treatment scientifically based on an understanding of these problems.

Case I evolves around the deep-rooted Jewish psychology of "Kadish," a mourning prayer recited by a son after the departed parents. This prayer does not mention a word about death or the departed. It is merely a recital of the glorification of God. But so deeply rooted is this custom that to a Jew it means everything. In it he sees immortality. When he is gone his child—part of him—carries on his ideals. This psychology was briefly expressed by a wise Jew who was congratulated on his son's accomplishments in the academic world, whose fame was world-wide, but who was not strict in his religious observances. He replied, "Tzu a 'Kadish' kumt is nicht." (Still he is not what is meant by a "Kadish.") In Case I the woman risked her life, in spite of the advice of her physicians, and gave birth to a son. Here is a deep subconscious motivation that prompted her to take the risk. The son was given a professional education at more sacrifice; and then he married out of his faith—which made him a lost cause. If he had died during the war, it would have been a death for the glory of the country and the flag. Psychologically there was death without glory and instead of glory there was insult upon injury. To this mother it was a severe psychic trauma which revealed itself in the various somatic manifestations. This problem was brought to the surface, but it was difficult to integrate it in her conscious mind and to make her see that he was still alive and still part of her. Her rational processes were blocked by intense emotions. One could not reason with her. She would not listen to reason. In order to focus her attention, it was decided to tell her an innocent story—a kind of psychological detour as an approach to the very core of her rational processes—a superficially innocent story in which a great Rabbi, who in the eyes of the religious Jew is final authority in questions of law, plays the main role. In the Rabbi's case his own daughter leaves him and his environment when she grows up. The story also reveals a man who, though converted, is not considered a total loss by the Rabbi. If such a man asks for mercy he still has a chance.

The second case deals with a keen business man who becomes a victim of a very serious

ailment, angina of effort. He is retired from his busy daily routine and every day reads the trends in the stock market quotations as well as the death column. One cause of death he understands—so-and-so died of heart attack. Death stares at him and he develops an intense fear of it until it finally becomes an obsession with him. This is not calm resignation to it. The fear is so intense that all words of encouragement rebound. The road to intellectual insight is blocked by fear. Therefore, an approach is made again by way of a detour—side-tracking his illness by focusing his mind on the romance of Palestine. The rejuvenation of a people is portrayed—the drama of Israel for many centuries. It is a fascinating story to listen to. While still absorbed in the story a suggestion is made to him, that in five years he should go to Palestine and see this drama staged before his own eyes. He swallowed a smooth pill without effort. The pill was further sugar-coated by appealing to his devotion to his wife, as such an experience would appeal to her religious instincts. As time goes on and nothing happens to him, his fear lessens and he has a goal, a five-year-plan. To lessen the emotional shock of the sudden death of a son to a person suffering from angina pectoris, another procedure was used. The tragedy of the accident was somewhat side-tracked by an appeal to him to cooperate in saving his wife, to whom he was devoted. She, on the other hand, was taken into confidence to spare the life of her husband. To further dampen the tragic jolt, I advised the family to have the funeral in the West and the temporary burial there. Meanwhile, the father would fulfill his duties by merely reciting the "Kadish" at home on the day of the funeral. Such was the case.

The third case illustrates an old Jew whose mental picture of an old Polish hospital made him dread all hospitals and close his eyes and ears to any suggestions about going to the hospital, no matter how urgent the case may be. Fear blocked the pathway to intellectual insight. It was opened by a back door. By telling the family to respect his decisions no matter what they may be, as it is the duty of children to respect their parents, I gained his confidence and made him feel that the physician was an ally to his cause. I then asked him to do me a favor and tell the "Angel" that the doctor had done his duty and so the cause of death, if and when it came, would be his own. This is not merely frightening a patient with death, for his fear of hospitals was even greater; but it has a more profound appeal. The religious Jew is taught that to preserve his own life is as sacred as saving the lives of others. Committing suicide is looked upon as a mortal sin and such a person is buried near the wall of the cemetery. Thus, the dreadful mental pictures associated with hospitals vanished from his mind and he listened to advice.

The last case illustrates a mother who lost her only daughter because of a postoperative accident and who focuses the guilt on the innocent surgeon. Attempts to explain this accident merely made the physician a partner to the guilt, one who belonged to the same "union." She felt that the physician knew the facts as she imagined them to be, but that he was merely concealing them. An iron curtain blocks all appeals to reason. Intellectual insight is gained by a sudden interruption of the discussion by the recital of an innocent prayer, one which is repeated daily in the orthodox mode of living. This prayer made clear to her the logic of embolic phenomena. The authors of prayers cannot be accused of union membership. Psychologically there is, indeed, power in prayer.

To conclude, these four cases serve as a reminder of a definite law about human mental behavior. Just as on the physical side, in order to understand the functions of an organ we study its evolutionary development, embryology as well as the anatomical and microscopic studies, so in a mental pattern of a patient we must not only delve into the complete history of a person but in the cultural background of that particular race from which he sprung. Every person carries within him mental seeds of culture from numerous generations that anteceded him. In the words of a poet and psychologist who said:

> "Our deeds still travel with us from afar,
> And what we have been makes us what we are."

SUMMARY

An attempt is made in this paper to show how significant it is for the physician to know the cultural background of the patient in order to diagnose and treat certain beliefs and traditions which are deeply rooted in the cultural history of a people. When these beliefs are challenged, in some instances psychosomatic problems develop. These four cases deal with Jewish patients whose people have a long, continuous cultural history with traditions and beliefs going back for many centuries. In such cases it is well that the physician himself be acquainted and familiar with the evolution of those ideas. In this way he can reveal their complexities and bring their problems to the surface.

Not only will a knowledge of the cultural background of the patient help in the diagnosis of the disease but it can also be made to act as a therapeutic agent.

Received April 2, 1951.

CRITIQUES AND NOTES

SELECTIVE PERCEPTION IN THE INTERPRETATION OF SYMBOLS

HAROLD H. MOSAK AND FREDERICK J. TODD

Veterans Administration Mental Hygiene Unit, Denver, Colorado [1]

In the development of psychoanalytic theory much has been made of analogy and allusion to validate or to illustrate certain Freudian hypotheses. The use of allusion finds prominent expression in the interpretation of symbols, e.g., fire is often treated as a symbol of male sexuality, and to confirm this interpretation, allusion is frequently made to the Prometheus myth. Although Freud in his early writings formulated a system which included fixed symbols (1), he later modified his position so that greater use was made of the patient's personal associations in the interpretation of symbols. Nevertheless, a perusal of psychoanalytic writings would reveal that symbols are often interpreted as having fixed meanings independent of the personality of the person who expresses these symbols—in dreams, in literature, in art, or in advertising. When such interpretations are made, it is our hypothesis that a process of selective perception occurs and that the symbol interpreter responds only to those elements which confirm his biases and neglects or rejects those aspects which are counter to his own values.[2] That such a process occurs has been experimentally demonstrated by Postman, Bruner, and McGinnies (2). In order to point out how this process operates we cite the following illustration:

Recently at a case conference a psychoanalyst, in discussing the case, referred to the automobile as a fixed masculine symbol. To lend further emphasis to this interpretation he alluded to phrases used in automobile advertisements. The writers consequently examined current automobile advertisements describing thirteen makes of automobiles. In support of the analyst's interpretation we found such statements as the following:

"power-packed with plenty of zoom"
"double rigid"
"the abundant performance of Fireball valve-in-head power"
"blazing get-away—entirely new command of the road at any speed"
"stronger, *twice* as rigid, lasts longer"

However, further examination of these advertisements demonstrated that the same sort of validity could be attributed to the interpretation of the automobile as a feminine symbol for we find such language as the following:

"rich . . . in the spaciousness of its interior—roominess"
"more value in beauty and comfort . . . more value in safety protection"
"leads in those things that make driving *easier, safer, more comfortable*"
"has more room in it than any other standard 6-passenger sedan on the market."
"when you find how much room and restful comfort [in the] spacious interior . . ."
"with these bodies . . . you'll enjoy a far safer, quieter, more comfortable ride."
". . . gives you 'Fashion Car' styling. From the sweep of its lines and the soundness of its coachwork . . . to the last detail of its 'jewel box' interior . . . it's America's best dressed car."
". . . started with a spacious interior and practically built the rest of the car around it! Here is room 'to roam around in'—plenty of room to let the human body take the positions that are *natural* and therefore *comfortable*."

Finally one advertisement which urges one to "feel the surging power of the high compression 'Get-Away' engine!" follows this with the contradictory statement, "see how smoothly *she* rolls along . . . and see how wonderfully easy *she* is to handle" (italics ours).

On the basis of this limited survey it seems apparent that an equally valid case for attributing feminine qualities to the automobile as a symbol could be formulated. Examination of the cited phrases would indicate that advertisement writers seem to stress both masculine and feminine attributes in exploiting their wares, masculine traits including power, rigidity, and size, while feminine attributes include roominess, comfort, protection, and glamour. From this, several conclusions are possible. It would seem that symbols cannot always, if indeed ever, be interpreted independently of the individuals employing the symbols. It is altogether possible, moreover, in this specific illustration that the advertising phraseology employed is a reflection of the individual advertisement writer's unconscious processes rather than a universal expression. Possibly because the advertiser must appeal to both men and women, he unconsciously plays up both aspects. This would also suggest that the private meaning that the automobile has for the individual is important. For some it may represent a symbol of power,

[1] Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the authors are the result of their own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

[2] We can discover this selective process in the ordinary use of proverbs where to believe "Absence makes the heart grow fonder" contradicts "Out of sight, out of mind." Those who accept "Haste makes waste," of course, neglect the attitude that "He who hesitates is lost."

prestige, wealth, or beauty, in which case an interpretation along Adlerian lines might equally "explain" the symbol. It also becomes evident that caution may be exercised in using allusory material for validation of symbolic interpretation since selective perception may seriously distort these interpretations.

REFERENCES

1. FREUD, S. *The interpretation of dreams.* New York: Carlton House, 1950.
2. POSTMAN, L., BRUNER, J., & MCGINNIES, E. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, **2**, 142–154.

Received March 23, 1951.

THE PSYCHOLOGY OF RELIGION: SOME SUGGESTED AREAS OF RESEARCH OF SIGNIFICANCE TO PSYCHOLOGY

W. EDGAR GREGORY
College of the Pacific

HALF a century ago the psychology of religion was flourishing and all kinds of expectations were held as to its future. Starbuck conducted his researches on the conversion experience (8). Hall dealt at length with the relation of adolescence to religious experience (1). James wrote his *Varieties of Religious Experience* (4). There was even a *Journal of Religious Psychology* from 1904 to 1911. Seward Hiltner has recently given a good survey of this development in an article in the *Crozer Quarterly* (2).

But about the time of the first World War much of this interest subsided. What interest there has been since then has been largely in the adapting of the practices of clinical psychology to the uses of the ministry. Clinical training for ministers has become an amazingly popular field, but there is almost no material available using the techniques of social psychology, the psychology of personality, etc., in the study of religion. What has been done has a mediocre character—largely the rehashing of psychological principles by someone whose primary training is theological and who is merely grafting psychological theory onto his basic discipline.

Yet from the standpoint of psychology as a science, religion offers untold possibilities for research —research of more general value than simply the understanding of religion. This paper does not pretend to be exhaustive, but only to point out a few of the major areas in which such research can have general value.

THE SECT PERSONALITY

One of our major problems today is to understand the personality attracted to such militant minority groups as the Communist party, the Silver Shirts, the Christian Front, the Ku Klux Klan, etc. Religious minority groups are for the most part much less secretive than these political groups. Records should be much more available for study. (Of course, the problem of rapport is a vital one. Most minority groups tend to be extremely "in-group" conscious and very suspicious of the outsider.) It should be possible, however, to establish something of the dynamics of personality which leads an individual to identify himself with Jehovah's Witnesses, a sect that considers all religious groups agencies of Satan and all government as designed to oppress the saved. Here is probably as extreme a group as one can find, in the sense that it stands completely apart from the larger society and considers everything associated with that larger society as evil. In spite of this, however, there is little thought of their being traitors in the political sense and no necessity for the organization to operate in secret as it is for the K.K.K. and the Communist party.

The Christian Science church offers another good chance for study. Why does this highly organized group of upper middle-class people (even more female in its composition than the other churches of the country) have such a hold on its members? Does its pseudo-intellectuality have anything in common with the radical "intelligentsia"? To be sure, the Christian Science church is highly resistant to any attempt to explore its records—even opposing a federal religious census—but individual Christian Scientists are quite willing, even anxious, to discuss their beliefs and practices with sympathetic outsiders.

Then there is the whole group of Pentecostal, Bible, etc., churches which have grown up outside the ordinary religious community, gaining their strength from the disfranchised, the transient, the

homeless. They, too, tend to be highly critical of the more respectable portions of society. Is it possible that the social isolate has here found a means by which his self-respect can be maintained (or restored)? A study of the sectarian personality might here be invaluable.

The question might well be put: What is the appeal of these deviant groups to the individual personality? Is it similar to the appeal of deviant political groups? What type of personality responds to this appeal? Does it have anything in common with the personality attracted by deviant political groups?

CULTURAL ACCOMMODATION

Interest is often expressed in the tendency of institutions to perpetuate themselves after the goal for which they were organized has ceased to exist. Religious sects offer something of the same phenomenon. The group which was sectarian (withdrawn from society) a generation ago, tends now to take on more of the role of a participating social group. What have been the changes in the personality needs of its membership to produce this change? Does it now attract a different type of personality? Has it ceased to work toward the objectives originally set up? In other words, what are the internal dynamics of the group that make this shift possible? Open-minded leaders of such groups welcome sympathetic interpretation of what is happening to their fellowship and what is likely to happen in the future. But this knowledge is of more than sectarian interest. It is of value for understanding all our institutions. It is obvious that a changing constituency will change the emphases, but we need to know vastly more of the detail involved.

INTERGROUP ACCOMMODATION

The social psychologist is deeply interested in the ways in which independent groups (particularly nations) come to adapt to each other. Actually in the "council" or "ecumenical" movement he has an experiment ready set for him. Here he can see the dynamics of adjustment between sovereign groups in actual operation—and there is a long and carefully documented history of the process. Here one sees the careful and eager search for formulas on which organizations of the most diverse history and structure can agree. Here he sees the pressures making for accommodation even when the official philosophies refuse to countenance it (as when the Eastern Orthodox churches join the World Council of Churches, although their official doctrine is that the other churches are not true churches, and justify it because they will here be "witnessing" to these other churches). Here he sees powerful denominational leaders giving fervent verbal support to the ecumenical cause but balking at the details which will make it really work. Here he finds aggressive officials of denominations officially forbidden to participate in such councils turning to the leaders of the councils for technical information and assistance. He sees the pressure on behalf of financially hard-pressed national and international leaders for pooling their efforts, and he sees sovereign denominations trying to come to grips over the problem of "colonies" (new churches in new communities and unfruitful churches in dying communities). He sees the confusion when a highly centralized church like the Methodist and a completely decentralized one like the Baptist try to understand and cope with each other's policies.

The study of intergroup dynamics as shown in the "council" and "ecumenical" movement can furnish a wealth of material for understanding the problem of rival unions, of rival states, and of rival nations.

UNIONS OF ORGANIZATIONS

The church union movement has a lot of material for the psychologist, also. The story of the "long road" to Methodist union (6) should furnish a good deal of material on the forces at work to perpetuate independent organizations. The story of the union of the Congregational and Christian churches is a totally different one in which a small, virtually bankrupt fellowship was absorbed into a larger, more affluent one, but still kept some of its sectarian qualities even in the union. The bitter fight going on now within the Congregational Christian churches—in the attempt to forestall union with the Evangelical and Reformed Church (itself a recent union)—should throw a good deal of light on the factors making for recalcitrant stands of minorities on the question of organic unions. (In this case one should know that one of the bitterest opponents of the union is the son of a man who was tried for heresy by such a body as the proposed United Church may turn out to be and who has a great fear of centralization: that another man has for some years been in charge of one of the richest religious endowment funds in the country and foresees the likelihood of having to spread it twice as far as before—he claims to fear its complete loss in the courts, etc.)

GROUP STRUCTURE

We often wonder at people turning to authoritarian governments and psychologically we try to understand this preference. Actually, within our own culture we have the most diverse forms of political structure available for study in the various ecclesiastical structures. The Roman Catholic church is completely centralized and has been able

to discourage, apparently quite voluntarily on the part of its members, any serious criticism from within the ranks. The Methodist hierarchical structure is almost completely nondemocratic. Once a bishop has been elected he is pretty well beyond the control of his group in every way. Even his salary and expense account are paid from national funds and not from local conference funds. He can't even be reached through the purse strings! Yet the average Methodist minister would be the last to think of his organization as undemocratic. He has learned to adapt to it and to manipulate it, and he thinks of it primarily as an institution for maintaining his security. If he gets in trouble with the local congregation he will not suffer; the bishop will simply move him to another post. Because of this knowledge of the security that is theirs, the Methodist ministry is thus able to speak out in criticism of local conditions in a way that the ministry of almost no other denomination can. Of course, it is equally true that one who departs too markedly from the desires of the bishop will find himself continually relegated to the "back woods"—but most of them seldom think of that. On the other hand, one finds a completely nonclerical structure in the Mormon church with no specialized training, ordination, or any of the usual requirements of a distinctive ministry laid upon its leaders. Any layman can get elected to local office—even to the First Presidency. But the decisions of the First Presidency are usually pretty binding and not to be violated thoughtlessly. Of course, the First Presidency (and bishops in other denominations) is fully aware of the state of mind of the church before new policies are laid down, but knowledge of the ways in which this awareness is obtained and of the way the membership is sold on a new policy will well repay psychological study.

(Incidentally, the Mormons offer an interesting field of study in the completeness with which the church dominates the lives of its members. This church is successful in expansion only when it establishes a community that furnishes means of satisfying all the needs of its members—cooperatives, recreational organizations, etc. The psychologist might well study some of the mechanisms by which this community solidarity is maintained.)

INDIVIDUAL EXPERIENCE

The conversion experience offered the early psychologists a fruitful field for investigation. One might well ask the question as to why the conversion experience is no longer such a significant fact in the lives of individuals. What cultural changes have produced this change in individual need?

The mystical experience, although investigated in some detail by Leuba (5), Pratt (7), and Hughes (3), might well be reinvestigated in the light of more recent findings. What distinguishes the mystical experience from abnormal visions and illusions? Or is there any distinction? Whatever the answer, the study should be rewarding in understanding abnormality (whether psychotic or not).

THE PSYCHOLOGY OF BELIEFS AND DOCTRINES

Finally, one needs to ask the question of all religious beliefs and doctrines: What personal need did this doctrine originally strive to meet? Does this need still exist? Does the doctrine still meet it? Are there other more effective ways of meeting it? The answers to all these will be of significance to psychology in general.

Social psychology, the psychology of personality, differential psychology, and abnormal psychology all stand to widen their horizons by including religious phenomena within their scope. It seems a little ridiculous that we who have ventured into so many forbidden areas of life should hesitate to study objectively this vast institutional field where there is waiting such a richness of material still unexamined. The above suggestions but skim the crock. It might well repay one of the foundations to put someone (or a committee) at the task of exploring in more detail the possibilities for psychological research in this field.

REFERENCES

1. HALL, G. S. *Adolescence: its psychology and its relations to physiology, anthropology, sociology, sex, crime, religion, and education.* New York: Appleton, 1904.
2. HILTNER, S. The psychological understanding of religion. *Crozer Quarterly,* January, 1947.
3. HUGHES, T. H. *The philosophic basis of mysticism.* Edinburgh: T. & T. Clark, 1937.
4. JAMES, W. *The varieties of religious experience.* New York: Longmans, Green, 1902.
5. LEUBA, J. H. *The psychology of religious mysticism.* New York: Harcourt, Brace, 1925.
6. MOORE, J. M. *The long road to Methodist union.* New York & Nashville: Abingdon-Cokesbury, 1943.
7. PRATT, J. B. *The religious consciousness.* New York: Macmillan, 1920.
8. STARBUCK, E. D. *The psychology of religion.* New York: Scribner's, 1903.

Received January 22, 1951.

VOLUNTEER-ERROR IN THE KINSEY STUDY[1]

A. H. MASLOW AND JAMES M. SAKODA
Brooklyn College

WHEN subjects are selected by any procedure other than random sampling, the possibility of systematic bias in selection must be considered. This is especially true of studies in which subjects are obtained through volunteering procedures, as is done in so many psychological researches (1). This danger can be avoided if it can be demonstrated that the use of volunteers does not introduce systematic bias into the results of the study. This test of the presence of bias or, as we shall call it, of *volunteer-error*, is dependent upon the availability of data on a variable which is correlated with the experimental variable and which is available for an unselected group including both volunteers and nonvolunteers.

One such variable is the level of self-esteem which was found to be highly correlated with various kinds of sexual behaviors and attitudes. In 1937 a paper-and-pencil test of self-esteem level was developed which could be used for research with large numbers of subjects. The validity coefficient of this test was .91 and the reliability coefficient, .90. It was reported in 1940 and published in 1942 by the Stanford University Press as the "Social Personality Inventory for College Women" (3).

The possibility of the use of the self-esteem test to study volunteer-error in sex studies was discovered in a study of sexual habits of women which Maslow reported in 1942 (4). He had used volunteers for the study, and stumbled across the disquieting fact that: (a) the volunteers were predominantly high in self-esteem (i.e., self-confident, sure of themselves, forward), and (b) those high in self-esteem score differed considerably from those low in self-esteem score in their sexual behavior.

It was found that about 90 per cent of the subjects were in the middle and high self-esteem groups, and only about 10 per cent in the low self-esteem group. An extremely high proportion of the low self-esteem subjects were virgins, were nonmasturbators, and were low in unconventional sexual attitudes, especially in oral-genital behavior. In contrast, many of the high self-esteem volunteers were nonvirgins, masturbators, and high in unconventional sexual attitudes. The self-esteem score correlated with unconventional sex attitude .71, with virginity —.66, and with masturbation .41. To correct for the volunteer-error a special effort was made to obtain subjects who were low in self-esteem score. In addition, a simple statistical technique for correcting for volunteer-error was utilized and described in this paper.

[1] We wish here to acknowledge the help given by Charles Fulweiler and Ed Davies, graduate students at the University of California.

Design of the Experiment

These considerations were urged upon Dr. Kinsey about 5 or 6 years ago, since his now well-known research at that time rested almost entirely on data derived from volunteers. A crucial experiment was jointly designed and executed by Dr. Kinsey and Maslow to employ the self-esteem test to evaluate the extent of the volunteer-error in Dr. Kinsey's study and to check on the correlations between self-esteem and sexual behavior which had been found previously. Drs. Kinsey and Pomeroy set up an office near the Brooklyn College campus, and Maslow made a strong appeal for volunteers in five of his classes in the psychology department, in which all students had already been tested for level of self-esteem by use of the Maslow Social Personality Inventory. It was then possible to compare these scores for the students who had actually furnished sex histories to Dr. Kinsey or to Dr. Pomeroy and those who had refused. Since the appeals to volunteer were made in class, the external pressure on volunteers and nonvolunteers was identical.

The students of all of these five classes had for other reasons also filled out Maslow's S-I Inventory (5), a measure of feelings of security-insecurity, and this was included in the analysis along with the dominance scores as a control, inasmuch as there was no reason to expect correlations between security and sex, or between security and volunteering.

Results

The difference between volunteers and nonvolunteers in the security-insecurity score was small and statistically insignificant, as expected, and can be omitted from further consideration. This excludes the possibility that volunteers were more neurotic than nonvolunteers.

The frequency distributions of volunteers and nonvolunteers on self-esteem score for five groups are given in Table 1. Data on an additional 64 volunteers not in these five classes are omitted here because they do not change the results and comparable data for nonvolunteers are not available. Some of the pertinent data for five classes comparing volunteers with nonvolunteers on self-esteem score are given in Table 2. The percentage of volunteers varies from 32.0 per cent in Group I to 56.7 per cent in Group II. For the groups combined

TABLE 1

FREQUENCY DISTRIBUTION OF SELF-ESTEEM SCORES FOR FIVE GROUPS OF VOLUNTEERS AND NONVOLUNTEERS

SELF-ESTEEM SCORES	GROUPS									
	VOLUNTEERS					NONVOLUNTEERS				
	I	II	III	IV	V	I	II	III	IV	V
143 – 157		2								
128 – 142		1								
113 – 127				1						
98 – 112		1		1	1		1	1		
83 – 97	2					1	1	1		2
68 – 82				1	3	1				1
53 – 67	1		2	1		2	3	1		1
38 – 52	2		1	1	3	2	2	1		
23 – 37			1	4	1	2		1	3	2
8 – 22	1	2	4	1	1	1	2	2	3	2
−7 – 7		4	1	2	2	2		1		3
−22 – −8		1	2		2	2	3	5	2	2
−37 – −23		4		1	2			3	6	2
−52 – −38	1			1			1	4		1
−67 – −53		1				2			1	3
−82 – −68						2				1
−97 – −83	1									1
−112 – −98		1								

the percentage of volunteers is 43.0 per cent. The variability among volunteers ranges from a standard deviation of 25.9 in Group III to 75.0 in Group II. Among the nonvolunteers the variability ranges from 45.3 in Group III to 89.2 in Group IV. For the groups combined there is more variability among volunteers (SD of 53.3) than among nonvolunteers (SD of 47.2).

The mean self-esteem scores for volunteers and nonvolunteers are shown in Table 3. Except for Group II, the volunteer groups consistently have a higher mean self-esteem score than the nonvolunteer groups. For all groups combined the mean self-esteem score of volunteers is 24.8 and for nonvolunteers 5.1, a difference of 19.7. The standard deviation for the combined volunteering and nonvolunteering group is 50.2, and the difference, which is .39 as large as this standard deviation, cannot be said to be an excessively large one. Figure 1 shows that there is considerable overlap between the distribution of self-esteem scores for the two groups. The difference, however, is a noticeable one, and one which is statistically significant at the .02 level of confidence by the *t*-test of combined groups and at the .01–.05 level by the *F*-test of five classes. The difference between volunteers and nonvolunteers in self-esteem score was not quite as large as the one found by the senior author in his study. This finding, however, upholds the hypothesis that on the average volunteers are higher in self-esteem score than nonvolunteers.

Analysis of the data showed that Group II, whose volunteers were distinguished by having a lower self-esteem score than nonvolunteers, was also distinguished by a number of differences from the other four classes. This class was a class in abnor-

TABLE 2

COMPARISON OF VOLUNTEERS (V) AND NONVOLUNTEERS (Nv) IN NUMBER, PERCENTAGE OF TOTAL, AND STANDARD DEVIATION

		GROUPS					
		I	II	III	IV	V	COMBINED
Number	V	8	17	11	14	15	65
	Nv	17	13	20	15	21	86
Percentage of Total	V	32.0	56.7	35.5	48.3	41.7	43.0
	Nv	68.0	43.3	64.5	51.7	58.4	57.0
SD	V	62.0	75.0	25.9	48.1	43.1	53.3
	Nv	53.6	46.6	45.3	89.2	52.0	47.2

TABLE 3

COMPARISON OF VOLUNTEERS AND NONVOLUNTEERS IN MEAN SELF-ESTEEM SCORE

	GROUPS					
	I	II	III	IV	V	COMBINED
Volunteers	24.9	17.4	19.5	32.6	29.7	24.8
Nonvolunteers	9.6	31.2	2.2	−9.0	−2.0	5.1

mal psychology with a high proportion of upper classmen, whereas the other four classes were classes in normal personality, composed mostly of sophomores and juniors. In the abnormal psychology class there were more students with extremely high or extremely low scores in self-esteem than in the normal personality classes, and the extreme individuals were volunteers rather than nonvolunteers. These characteristics indicate the subtle fluctuation in selection that can be introduced by the use of volunteer subjects. Clearly more research is desirable in this area.

The second part of the experimental design which was to test the correlation between self-esteem score and sexual behavior in Dr. Kinsey's subjects has not been reported to date by Dr. Kinsey. We are, therefore, unable to report the extent of correlation between self-esteem score and sexual behavior for Dr. Kinsey's Brooklyn College subjects. Unsatisfactory as this is, we must fall back on Maslow's previous study and assume that the correlations found there apply also to the subjects of the present study.

DISCUSSION

Perhaps the important conclusion to be drawn from this experiment is a methodological one—namely, that a technique is available to test for volunteer-error. Because of the intimate and unconventional nature of sex studies, it is not feasible to abandon the use of volunteer subjects or voluntary return of questionnaires. We believe, however, that where such volunteering procedures are employed a test for volunteer-error should accompany it before the results are published as representative of the total population of volunteers and nonvolunteers. It is probable that self-esteem score can be used as a test variable to check volunteer-error, not only in the study of sex, but also in studies of other unconventional forms of behavior.

The problem of controlling bias in selection when strictly random samples are not employed is an old one in surveys. The general procedure developed to meet this problem has been to stratify the population by a number of variables which are correlated with the behavior to be studied. Dr. Kinsey in his sex study has taken similar precautions by stratifying his population by such factors as age, sex, religion, occupation, education. This study indicates

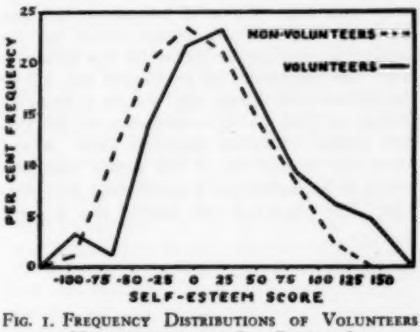

FIG. 1. FREQUENCY DISTRIBUTIONS OF VOLUNTEERS AND NONVOLUNTEERS IN SELF-ESTEEM SCORE

that such stratification on the basis of objective indices alone, largely sociological in nature, may not be sufficient to eliminate volunteer-error, at least in the case of sexual studies. There is need to investigate psychological variables related to volunteering to discover and control bias within each stratified group.

To check the representativeness of his samples Dr. Kinsey has been getting some "100 per cent sample groups," in which every member of the group could be used as subject. Such a step will help in correcting for volunteer-error within a particular group, but not for volunteer-error for use of groups which volunteer 100 per cent as against groups which are unwilling to volunteer 100 per cent. Here again a test for volunteer-error is desirable.

The second conclusion relates more directly to sex studies in general and to Dr. Kinsey's study in particular. If our inference is correct, the bias introduced into a sex study by the use of volunteers is, in general, in the direction of inflating the percentage reporting unconventional or disapproved sexual behavior—such as masturbation, oral sexuality, petting to climax, premarital and extramarital intercourse, etc. The more timid and retiring individuals, evidently, are apt to be privately, as well as socially, conforming. They are likely, it seems, to refrain from volunteering for sex studies in which they are asked embarrassing questions. The present study would lead us to conclude that the percentages reported are probably inflated and that they should

be discounted to some extent for volunteer-error until reexamined.

This criticism is not to be interpreted as an attack upon the central principles of the Kinsey study. On the contrary, we heartily approve of the basic procedures, but believe the study to be so important that no possibility of refining the techniques should be overlooked.

SUMMARY

A previous study found that volunteers for a sex study were high in level of self-esteem, and that this was correlated with unconventional sex behavior. A technique for testing for this volunteer-error was suggested. An experiment was jointly planned with Drs. Kinsey and Pomeroy at Brooklyn College to check on the introduction of bias into their samples by use of volunteers alone. It was found that in four out of five groups volunteers tended to be higher in self-esteem than nonvolunteers. The conclusion was reached that a check should be made on volunteer-error in studies of unconventional behavior in which volunteer subjects are employed, and that reported data from sex studies using volunteers need to be discounted for volunteer-error.

REFERENCES

1. BROWER, D. The role of incentive in psychological research. *J. gen. Psychol.*, 1948, **39**, 145–147.
2. KINSEY, A. C., POMEROY, W. B., & MARTIN, C. E. *Sexual behavior in the human male.* Philadelphia: W. B. Saunders, 1948.
3. MASLOW, A. H. A test for dominance-feeling (self-esteem). *J. soc. Psychol.*, 1940, **12**, 255–270. Published 1942 by the Stanford University Press as *Social Personality Inventory for College Women.*
4. MASLOW, A. H. Self-esteem (dominance-feeling) and sexuality in women. *J. soc. Psychol.*, 1942, **16**, 259–294.
5. MASLOW, A. H., BIRSH, E., STEIN, M., & HONIGMANN, I. A clinically derived test for measuring psychological security-insecurity. *J. gen. Psychol.*, 1945, **33**, 21–41.

Received January 24, 1951.

TWO RORSCHACH SCORING TECHNIQUES: THE INSPECTION TECHNIQUE AND THE BASIC RORSCHACH[1]

ANNE ROE
New York City

THE two major techniques advanced for handling of Rorschach data in more quantifiable and standardized manner are the Munroe Inspection Technique, which applies either to group or individual administration (2), and the Buhler-Lefever Basic Rorschach, which applies only to individual administration (1). The two methods will not be described here, but they have essentially the same aim, and the basic approach is similar. In the course of a study of eminent scientists, 61 Rorschachs were administered individually. All have been scored by both techniques, and a brief comparison of the results on the two systems is presented here.

It should be emphasized that the Buhler-Lefever administration differs somewhat from that employed in this study, and these authors point out correctly that their method is strictly applicable only when their procedure of administration and scoring is followed. The differences in scoring between the two methods are minimal and are easily handled. I have altered the scoring for the Basic Rorschach in accordance with their Manual. A more difficult matter is the difference in administration. The major problem here is that I put no limits, either way, on total number of responses. The Basic Rorschach administration, on the other hand, asks specifically for three to five answers to each card, and the subject is encouraged to give three and discouraged from giving more than five, although permitted to do so. Analysis of my results showed that for protocols of up to 80 or 90 responses, there was so little difference in the Basic Rorschach scoring that it was hardly worth while to rescore on the basis of the first six responses on each card. For protocols longer than 70, however, I have retabulated the scores on the basis of not more than six responses per card. No corrections can be made for the shorter protocols.

Figure 1 gives scatter diagrams for each scoring system against total number of responses. (These are different in the two diagrams, because of the rescoring for the Basic Rorschach.) Correlations are significant in both instances, being +.314 for the Inspection Technique and +.477 for the Basic Rorschach. This would mean that with the Inspection Technique there is some tendency for the longer protocols to be associated with poorer adjustment, but on the Basic Rorschach there is an even stronger tendency for the longer protocols to be associated with better adjustment. (The higher scores on Inspection Technique indicate poorer adjustment.) It is possible that both relationships are curvilinear. It would appear, however, that the sort of sliding scale of the Inspection Technique, although some correction is indicated, does take fairly adequate account of differences in total number of responses. In the case of the Basic Rorschach, it would seem that the technique should probably not be applied, unless the administrative technique of requesting a minimum of three responses per card has been used. It should be noted, of course, that this will not always get three responses; I am practically certain that with these subjects, at least four of the ones who gave short protocols would have been unable to add to them, and I doubt if many of the others would have gotten up to a total of thirty, although doubtless some of them would.

Figure 2 is a scatter diagram for the group on the two scoring systems. This correlation is −.505 (higher scores on the Inspection Technique indicate poorer adjustment). There are some striking discrepancies (e.g., the best-adjusted man by the Inspection Technique rates in Level III in the Basic Rorschach), and the correlation is low enough that it is clear the measures are far from interchangeable, although both purport to be measures of adjustment. The Basic Rorschach was standardized on a varied population, the Inspection Technique on a college group and doubtless much of the discrepancy is accounted for on these grounds. These comments are not intended as a criticism of either technique, but the whole point to such techniques is their general applicability. A detailed analysis of the discrepancies, which is beyond the scope of the present paper, should shed considerable light on the whole problem of criteria of adjustment, and would be a most useful service.

Something further should be said about the general implications of the scores on both of these techniques. These subjects are all men of very great attainment. Nevertheless, some of them obtain scores on one or both of these two techniques which indicate very severe maladjustment. It is clear that many of them do have serious emotional

[1] Part of a study financed by a grant from the Division of Research Grants and Fellowships of the National Institutes of Health of the United States Public Health Service.

problems, but it should be emphasized that they have been able to handle these successfully in one way or another. I have discussed this point more fully elsewhere. Here I should just like to point out some implications of these distributions.

The mean Inspection Technique score for this group is 10.8. Munroe gives no general cut-off

FIG. 1. SCATTER DIAGRAMS SHOWING RELATIONSHIPS BETWEEN TOTAL NUMBER OF RESPONSES AND INSPECTION TECHNIQUE SCORE, AND TOTAL NUMBER OF RESPONSES AND BASIC RORSCHACH SCORE. N=61.

points for levels of adjustment, but says these must be considered separately for any new group of subjects, and for any particular purpose. She does, however, state that a score of 10 or above would indicate, in her college group, some need of particular attention. Comparable scores are available for a group of successful artists who averaged 10.3. A group of 382 faculty members given the Group Rorschach averaged 10.1 on this measure. Inspection Technique scores for this group of eminent men range up to 19, so that if the measure has any validity, it would appear that a number of these subjects are quite seriously disturbed. They are even more definitely so labeled by the Basic Rorschach scores.

In the monograph on the Basic Rorschach four levels are described as representing "four degrees of reparable or irreparable loss of the personality's capacity to master conflict situations and to meet reality."

Level I "adequacy" is described as one in which there is control over goals and over emotionality, with less anxiety than on Level II but considerable tension. Level II "conflict" shows the characteristic neurotic conflict between tendencies toward immediate satisfaction and tendencies toward deferred satisfactions, with the peaks of uncontrolled emotionality, anxiety, insecurity, and unsatisfied instinctual

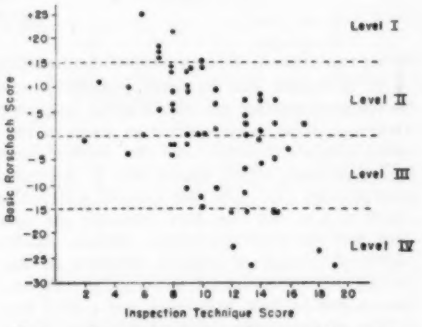

FIG. 2. SCATTER DIAGRAM SHOWING RELATION BETWEEN INSPECTION TECHNIQUE SCORE AND BASIC RORSCHACH SCORE. N=61.

drives. On Level III "defect" goal deferment is abandoned, drive rules in a stereotyped manner, and anxiety as well as emotional responsiveness is either repressed or deficient. In Level IV "reality loss" the problem no longer exists; all striving appears to be reduced below the conflict level; reality is beyond reach.

As can be seen, eight of these subjects would fall into Group IV and eighteen into Group III (assigning those with scores on the dividing line to the higher group).

All of this would strongly suggest that in dealing with normal subjects, the greatest care must be taken in interpretation from the Rorschach to probable behavior. The usual clinical standards do not apply. In such records as I have collected one is impressed repeatedly with the fact that the Rorschach may not give any hint of the actual life performance of the subject. It may indicate the presence of severe problems, and it must be said in justice to the Rorschach that the reality of these can usually be determined, but it very likely will give no indication of the fact that the subject has managed to cope very effectively with the problem. Knowing that he has coped, one may be able to tease out some information as to how he has done it, but even this is not always possible.

REFERENCES

1. BUHLER, CHARLOTTE, BUHLER, K., & LEFEVER, D. W. Development of the Basic Rorschach Score with Manual of Directions. 1948. Privately printed.
2. MUNROE, RUTH L. The Inspection Technique: A method of rapid evaluation of the Rorschach protocol. *Rorschach Res. Exch.*, 1944, **8**, 46–69.

Received May 6, 1951.

A NOTE ON THE VALIDITY OF SANFORD'S AUTHORITARIAN-EQUALITARIAN SCALE

JOAN EAGER AND M. BREWSTER SMITH

Vassar College

THE upshot of a varied and growing body of research testifies to the highly ramified importance of a psychological dimension identified at its negatively valued extreme as authoritarianism. Its purity as a dimension, to be sure, has not been demonstrated, and further clarification of its psychological nature remains to be achieved by factor analysis and scaling methods. Nevertheless, the relevance of some such personality variable has been established in areas as diverse as ethnic prejudice (1), cognitive rigidity (4), perceptual intolerance of ambiguity (3), and citizen orientation toward political leadership (5).

Further research with this variable must hinge to a considerable extent on the refinement of appropriate measuring instruments and their adaptation for use on a wider scale. Sanford and Older (6) have devised an eight-item Authoritarian-Equalitarian (A-E) Scale as a step in the latter direction. Their abbreviated measure, adapted from the longer "F-scale" of the California studies of prejudice (1) for use in door-to-door sampling surveys, was reported by its authors to have a repeat reliability of .78 and a split-half reliability of .58. On a small college sample they found a correlation with the F-scale of .67. Sanford presents considerable evidence showing that followers' scores on his brief scale are associated in a coherent and intelligible way with their orientation toward authority figures, presumably an important determinant of the forms of leadership that prevail in any group. But his data, drawn from interview and projective procedures that explored the psychological environment of the follower, provide no direct evidence for the validity of the scale so far as the overt behavior of either follower or leader is concerned.

Below we report some evidence, from a study (2) of a small number of appointed leaders in a natural setting, that permits tentative conclusions in regard to the relation of scores on Sanford's A-E scale to behavioral consequences. The weight that has been attributed to the authoritarian-equalitarian dimension, we believe, lends to these findings an interest beyond what might otherwise be warranted.

METHOD

Sanford's A-E scale was included in a questionnaire administered to the eleven counselors in an endowed "health camp" for underprivileged children in up-state New York, before the opening of the camp. The counselors were all college girls from different colleges. The greater number of them had little or no previous camp-counseling experience. The first month-long period was for boys, followed by a similar session for girls. About 60 children of each sex between the ages of 6 and 11 attended. With the camp administration laying its emphasis on improving the physical health of the children, there was little explicit attention by the directors to psychological problems of counselor-child relations. While counselor-led group activities were provided, facilities were not available for paying much notice to the special problems of individual children; boys or girls who showed any major disturbance were sent home. If we were to hazard a gross characterization of the prevailing social atmosphere, it would clearly be on the authoritarian side.

Responses to the seven items of the A-E scale found most discriminating by Sanford were used to dichotomize the counselor group into five relatively equalitarian and six relatively authoritarian counselors. The items were these:

1. Human nature being what it is, there must always be war and conflict.
2. The most important thing a child should learn is obedience to his parents.
3. A few strong leaders could make this country better than all the laws and talk.
4. Most people who don't get ahead just don't have enough will power.
5. Women should stay out of politics.
6. People sometimes say that an insult to your honor should not be forgotten.
7. People can be trusted.

A choice among six answers was provided for each item, arbitrarily scored from 6 to 1, respectively: agree very much, agree pretty much, agree a little, disagree a little, disagree pretty much, disagree very much. For item 7, the scoring was reversed. Scores for the "equalitarian" counselors ranged from 10 to 16, with a mean of 12.5, while the "authoritarian" counselors' scores ranged from 17 to 26, with a mean of 21.2.

The second source of data was a "Guess Who" questionnaire administered in individual interviews to 17 boys and 17 girls in the last two days of each camping month, to get at the children's perceptions of the counselors' behavior and characteristics, particularly in regard to authoritarian behavior. The authoritarian extreme, as it turned out, was over-represented in the items; had the items provided a more even gradation between the authoritarian and equalitarian poles, they should have comprised a more sensitive instrument. The subjects were selected by taking every third name from the cabin rosters, approximately equal numbers being drawn from each of the three cabins in which the children and counselors lived.

RESULTS

The results of the "Guess Who" questionnaires from the two camp sessions were analyzed by noting the number of "equalitarian" and of "au-

TABLE 1

THE RELATIONSHIP BETWEEN CAMPERS' PERCEPTIONS OF COUNSELOR BEHAVIOR AND COUNSELORS' RESPONSES TO SANFORD'S A-E SCALE

ITEM	NUMBER OF COUNSELORS MENTIONED: (a) "Equalitarian"	(b) "Authoritarian"	$\frac{a-b}{a+b}$
1. She has a quiet group that pays attention and has fun. (9)	20	9	38%
2. She always helps you when you need it in learning something new and seems very glad to do it. (3.5)	18	9	33%
3. She explains how to do something to your group and then leaves the rest up to you. (5)	17	9	31%
4.5 You don't want to do what your group is supposed to but she gives you a special job or makes you like it. (9)	11	7	22%
4.5 She asks you what you want to do at the beginning of a period and lets you do what you want. (2)	14	9	22%
6. She has a quiet group because they are afraid of her. (30)	15	10	20%
7.5 She tells you what to do, and won't let you do anything else. (28)	16	11	19%
7.5 She has a group that runs around and doesn't pay any attention to her. (12.5)	16	11	19%
9. She doesn't make you play when you don't want to, and she doesn't give you anything else to do. (12.5)	17	12	17%
10. You have fun with her and like to be with her. (3.5)	17	13	13%
11.5 When you don't want to play with your group she makes you or else punishes you. (24.5)	14	11	12%
11.5 You do what she asks you to because she usually wants to do what you want to do. (9)	14	11	12%
13. She talks to you like you are her friend. (1)	13	11	9%
14. She doesn't help you do or make things and then yells at you when you do them wrong or badly. (24.5)	12	11	4%
15. She hardly ever gets mad at anyone. (6.5)	14	14	0
16. She is always telling you not to do something. She is always saying stop that, don't do that, shut up, etc. (30)	16	17	−3%
17.5 She lets you pick what you want to do from two or more things. (6.5)	12	13	−4%
17.5 When you have done something wrong, she bawls you out in front of the whole group. (24.5)	11	12	−4%
19. You do pretty much as she asks or tells you to because you like her and don't want to cause her any trouble. (11)	13	15	−7%
20. When you have done something wrong, she hits you. (24.5)	9	14	−22%
21. You don't do what she asks you to because you don't like her. (15.5)	7	12	−26%
22. When you have done something wrong, she threatens to do something to you next time like sending you to ..... (name of camp directors). (24.5)	8	14	−27%
23. When you have done something wrong, she just tells you to stop it and not to do it again. (14)	9	16	−28%
24. She gets mad at every little thing. (19.5)	8	15	−30%
25. You don't do what she asks you to because you don't want to do what she usually asks. (18)	6	12	−33%
26. She just gets mad when someone has been bad. (15.5)	9	19	−36%
27. She talks to you like you were a baby. (21)	6	14	−40%
28. When you have done something wrong or been bad, she sends you to the front porch. (19.5)	5	12	−41%
29. You do what she asks you to because she will punish you if you don't. (30)	5	14	−47%
30. She talks to you like she is always mad at you. (24.5)	4	19	−65%
31. You don't do what she asks you to because you like to make her mad. (17)	2	13	−73%

thoritarian" counselors, as classified by A-E score, named for each item. The latter number was subtracted from the former, and the difference expressed as a percentage of the total number of responses given to the item. The items are listed in Table 1, and arranged in an order starting with those for which "equalitarians" were named with the greatest relative frequency.

DISCUSSION

These data in Table 1 are based on too few cases to warrant detailed attention to the individual items. What is relevant is the extent to which a coherent pattern emerges which distinguishes between the authoritarian- and equalitarian-scoring counselors as they were perceived by the campers. Casual inspection, indeed, indicates the presence of such a pattern, sharply contrasting, for example, the first and last five items when they are ordered as above. More objective evidence is available, however, that this apparent pattern actually conforms to what would be predicted on the assumption that A-E scores were reflected in counselor behavior. The 31 items, typed on cards, were rated on a five-point authoritarian-equalitarian scale by five psychology and sociology instructors. The items were then ranked according to the average rating assigned by the judges, a low rank indicating a relatively equalitarian average rating. These item ranks are given in Table 1 in the parentheses following each item. When a rank-difference correlation was computed between judges' rankings and rankings by "Guess Who" nomination, a rho of +.44 was obtained. In view of the brevity of the A-E scale, the small number of counselors, and the inadequacies of the "Guess Who" items (reflected in numerous tied ranks by judges' ratings and in a clustering of judgments near the authoritarian end of the scale), the magnitude of the correlation gives considerable evidence that the attitude dimension tapped by the scale is a valid predictor of actual counselor behavior.

Supporting evidence comes from a separate analysis of the "Guess Who" results from the boys' and girls' sessions. Since the A-E scale was administered prior to the boys' session and two full months elapsed before the second "Guess Who" test was given to the girl campers, one might expect a higher correlation between "Guess Who" results and judges' ratings for the boys than for the girls if, as additional data give us reason to believe, real changes occurred in A-E scores during the camp sessions. A further ground for predicting a higher correlation in the boys' group stems from the fact that the boys, being more unruly, may have been more effective in bringing out latent authoritarian trends in the counselors. In accordance with expectation, the higher correlation was found for the boys' group: a rho of +.53 as compared with +.33 for the girls.

While the scope of the present study warrants only tentative conclusions, evidence has been presented showing that in a camp for underprivileged children, the campers' impressions of counselors who scored high on the A-E scale differed coherently from their impressions of low-scoring counselors. These results support the validity of the scale for studies of leader behavior as well as for studies of follower orientation toward leadership.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. EAGER, J. A study of leadership. Unpublished B.A. Thesis, Department of Psychology, Vassar College, 1952.
3. FRENKEL-BRUNSWIK, ELSE. Personality theory and perception. In R. R. Blake and G. V. Ramsey (Eds.), *Perception: An approach to personality*. New York: Ronald, 1951, pp. 356–419.
4. ROKEACH, M. Generalized mental rigidity as a factor in ethnocentrism. *J. abnorm. soc. Psychol.* 1948, 43, 259–278.
5. SANFORD, F. H. *Authoritarianism and leadership: A study of the follower's orientation to authority*. Philadelphia: Institute for Research in Human Relations, 1950.
6. SANFORD, F. H., & OLDER, H. J. *A short authoritarian-equalitarian scale*. Philadelphia: Institute for Research in Human Relations, Report No. 6, Series A, June 1950. (Mimeo.)

Received February 20, 1952. Prior publication.

REVIEWS

EDITED BY STANLEY G. ESTES

ORGANIZATION AND PATHOLOGY OF THOUGHT: Selected Sources. Translation and Commentary by David Rapaport. New York: Columbia University Press, 1951. Pp. 786. $10.00.

This book differs in several important aspects from the typical "Readings in . . ." which review major fields of interest. It is unique because it deals with thinking, a neglected area in psychology. For instance, among the 5,000-odd entries in the 1948 *Psychological Abstracts*, Rapaport found only 50 classified under "Thinking," and most of these dealt with concept-formation tests. (This reviewer found "no place for Thinking" as a specialty in the last national roster of professional psychologists.) As a collection of papers, several of which are excerpts from larger works, Rapaport's book is moreover unique, because, out of the 27 published studies, at least 20 are for the first time made available in English. And it is further unique, because, again for the first time, the psychology of thinking has been explored in both its normal and pathological manifestations within one volume. Finally, Rapaport has included his own commentaries on each article, and a final chapter of his own. His special contribution lies also in the selection of these studies from a problem-centered point of view, aiming at the clarification of issues rather than a survey of information.

The choice of articles indicates the scope of this attempt. There is first a section on *Directed Thinking* with selections from Ach's "Determining Tendencies"; from Buehler's "On Thought Connections"; a study by Claparède on "Recognition and 'Me-ness'"; two early but very representative papers by Lewin, one on basic methodology, and one on "Intention, Will and Need"; and two reprintings from Piaget's work dealing with intellectual evolution from childhood to adult life, the first presenting a psychological and the other a more biological point of view. The second section is on *Symbolism*, wherein various experimental observations are reported on the spontaneous or induced experience of specific symbolic contents during different mental states, such as semi-waking, hypnosis, and suggested dreams. The investigators are Silberer, Schroetter, Roffenstein, Nachmansohn, and Betlheim and Hartmann, the last two mentioned dealing with patients suffering from Korsakow Psychosis. The next section, on *Motivation of Thinking*, contains Stekel's article on "The Polyphony of Thought"; three papers by Freud; an analysis of "The Psychology of Boredom" by Fenichel; and Heinz Hartmann's "Ego Psychology and the Problem of Adaptation," perhaps the most far-reaching reformulation of psychoanalysis in recent decades. A section fellows on *Fantasy Thinking* with two articles by Bleuler on autisitic thinking; excerpts from "The Psychology of Daydreams" by Varendonck; and from Kris's paper, "On Preconscious Mental Processes." The final section presents four studies on the *Pathology of Thinking*, each of which also contains illuminating data on normal thought processes. This is particularly striking in the two contributions of Schilder, "On the Development of Thoughts," and his article on the psychology and symptomatology of general paresis. Schilder's analysis of how a thought develops in its preparatory phases should revive interest in the "anticipations" of meaning and the problem of how thought originates within the general "realm" or "sphere" of intended meaning. The selections from Bleuler's well-known "The Basic Symptoms of Schizophrenia" include protocols and analyses of patients' thought-products and have lasting didactic value for the clinician. The report by Buerger-Prinz and Kaila, "On the Structure of the Amnesic Syndrome" with its evidence for a generalized thought impairment, remains a classic demonstration of the use of experimental techniques for the study of clinical problems.

The richness and psychological depth of this material cannot be conveyed in a review. All these studies reveal not only the multi-layer organization of thinking but also its ramified interlocking with the person as a whole, however conceived. The tough-minded empiricist will meet many a conclusion which, by vigorous experimental standards, is still open to question. Yet he may find these ingenious inquiries into really basic problems a rewarding experience. The originality of approach in the earlier "sources" outlasts their historical outdatedness, and may inspire standards for more productive working hypotheses that are both empirically testable and psychologically acute. The combination of theorizing, experimentation, and penetrating analyses of individual cases in this book demonstrates anew how careful examination of qualitative data may help the search for relevant hypotheses and quantification.

The extent and intensity of the editor's scholastic labor is evident from the fact that for 684 pages of selections there are, chiefly provided by Rapaport, 1865 footnotes and a bibliography

of 787 references. The amount of text in the footnotes consumes large portions of most of the pages in the book. For instance, on 107 pages the footnotes cover three-quarters and more of the page. On 206 pages they fill at least half the page, and on 221 pages, over a quarter page. Only 12 pages go "unfootnoted." The footnotes are of two kinds: those by the original authors, quite moderate in number and extent, and those by Rapaport. The latter, in turn, consist of summaries of omitted passages; interpretive comments; criticisms; references to, and quotations or excerpts from other authors, especially Freud; cross references to other chapters or to other footnotes throughout the book or to footnotes in other publications.

The volume of this exegetic production amounts to "a book within a book," written in smaller type. But it also places an immense burden on the reader; it switches his "attention cathexis" from practically every other sentence in the text to the associative by-play in the editor's footnotes and then back again to the original discourse. This breaks up the continuity of the reader's comprehension. The effect is more diffusing than focalizing and tends to blur in the reader's mind (or in his "trace system") the distinction between the author's original composition and the editor's variations on the theme —which often become autonomous creations. Though the commentary provides information of encyclopedic scope, there is, nevertheless, considerable unevenness which varies from esoteric items for the psychoanalytic sophisticate to the obvious. However, closer scrutiny reveals that much of the material in "the small-type volume" approaches redundancy in the exposition of leading ideas and in bibliographic details. Identical passages or quotes are unnecessarily repeated,[1] and documentary digressions occur.

The displayed "polyphony of thought" and quotational capriciousness in a book on Organization of Thought lead to questions regarding the policy of selections: what was the basis for deciding *when* to introduce commentary, which *kind*, what amount, and for what type of reader? For example, though historical minutia are dealt with in the psychoanalytic references, the section on *Directed Thought*, which covers a period from Ach (1905) to Piaget (1936) lacks pertinent historical background information. Findings by researchers such as Gibson, Postman, and Bruner, who published much later than the translated authors, are introduced into the footnotes for the papers of Ach, Buehler, Lewin, and others. These big temporal discrepancies are not made clear to the less-initiated reader who meets, in addition, such contemporary names undated save by bibliographic number.

After one has read the book's final chapter by Rapaport—without footnotes—one is again tempted to ask questions about the editorial form. Evidently, the preceding articles, together with the running commentary, are intended to build up a supportive base for the central thesis of the book. To a large extent, the conclusions of his own chapter are reformulations of his "lower-case" annotations. Consequently one asks why Rapaport found it necessary to tell his message in two ways: one, in the cumbersome indirection of some 1800 elaborative comments on other authors, and, two, in a direct way. The answer that this question misses the whole point of a scholarly documentation does not dispose of another and just as scholarly alternative: to present the translations intact in their original form, to let them first speak for themselves, and then to append *after* each article the relevant commentaries in connected discourse. In this way, the original thought-organization of both the translated authors and the translating editor might have gained more effectively the distinction which each deserves in every sense of the word.

The viewpoint which underlies this particular combination of original sources and commentary seems to consist of the following ideas.

First, psychoanalysis and so-called academic psychology are not basically incompatible. They may probe different levels of thought phenomena and thereby supplement each other. Examples of such different levels are: Buehler's relationship-consciousness and interthought-knowledge, Selz's anticipatory schemata, Lewin's quasi-needs and tension systems created by acts of intending, Ach's unconsciously-acting Einstellungen, Schilder's preconscious sphere of meaning, Freud's projective representations of drives in imagery and thought, and so on. Thus the investigation of thought processes in the experimental laboratory may be supplemented by their study in the natural laboratory of pathology, the clinician also turning experimenter. In addition, "dynamic psychology" and child psychology are increasingly accepting and influencing each other. The maturation levels of thinking, perception and ego-development, e.g., in the language of Piaget and Werner, and the motivational stages of personality development in the language of psychoanalysis, seem no longer totally unrelated. The two universes of discourse may be mutually translatable, notwithstanding the fact that each system conceptualizes different developmental aspects.

Second, a meeting ground may also be found

[1] For example, No. 24, p. 323 and No. 15, p. 343; No. 35, p. 412 and No. 2, p. 438; No. 46, p. 416, No. 3, p. 439, and No. 31, p. 372; No. 11, p. 478, and No. 13, p. 353; No. 3, p. 439, and No. 46, p. 417; No. 26, p. 658, and No. 201, p. 559.

in the growing recognition which psychoanalysis has been giving to the "conflict-free sphere of the ego" and "autonomous" ego functions. These terms of Hartmann's refer to adaptive properties of the organism such as intelligence, perception, and motility. Their development has a basis in biological maturation rather than in motivation stemming from instinctual conflicts. These capacities, together with other "neutral" functions, constitute the "conflict-free sphere of the ego." This represents a departure in Freudianism from the "solipsistic tendencies" of drive-determined behavior (p. 526) and may lead to a new rapprochement between psychoanalysis and academic psychology. The former no longer centers on the "subjective" factor, on thoughts' *projective* nature; and the latter no longer centers exclusively on the "objective factor," on thoughts' *adaptive* nature.

Third, this reorientation leads to the central problem of the book, namely, the relationship of thought to motivation. From the psychoanalytic standpoint, there cannot be a psychology of thinking without a psychology of motivation. Yet the new trend towards "neutral" ego-functions may broaden the motivational theory of thinking to include problem-solving activity and creative thought as more nearly separate from the "vicissitudes of instincts." Rapaport proposes in this framework that one should go beyond the analysis of thought contents, and study thinking in its formal aspect, under which thinking is to be viewed as a mediating function between reality-adapted behavior and personality motivation. He realizes that this program may broaden the traditional definition of thinking as conscious problem solving to include all of psychology. He counters this criticism by an interesting argument—namely, that thought processes must be an aspect of most psychological activities, conscious or unconscious. In turn, therefore, consciousness can be only an *aspect* of *certain* thought organizations, more specifically an aspect of those which undergo a specific distribution of "attention cathexis." This proposition is of great consequence for any theory of normal and abnormal thinking. It is the basic theme which links together the various treatises on thought organization in the book. Without Rapaport's own summary, however, this theme would remain implicit; it becomes explicit in his conceptual model of the relationship between drive and thought.

This model represents a most brilliant resumé and extension of Freudian theorizing. Thinking, consciousness, and motivation are placed within the framework of cathectic dynamics, i.e., energy distribution, and ego-structure. The ontogenesis of these is traced through successive phases of mental organization, e.g., the stages of primary and secondary thought, conceptual thinking, and different states of awareness. These developmental stages become *levels* in a functional hierarchy of drive and thought. With each new level, there arise new regulatory controls over the former together with new autonomous functions and motives of their own. Though impossible to render the steps of this reasoning faithfully within the confines of a review, they may be summarized as follows:

1) Primary process. Drive-needs are disequilibria in energy distribution manifested in restlessness. The ensuing tendency towards regaining equilibrium leads to a tension directed towards (drive) objects of valence. Equilibrium is restored through the drive-satisfying object which provides the condition for tension discharge in motor action etc. (the pleasure principle). The model of action and thought thus conceived is derived from need-gratification in the infant: First, hunger —restlessness—sucking action on the breast—tension discharge; second, absence of drive object—mounting drive tension—hallucinatory image of the object. Because of the *delay*, the drive discharges energy into the memory trace of past gratification and makes it conscious. The memory of breast sucking becomes "cathected" to perceptual, "hallucinatory" realness in an attempted short cut towards wish fulfillment and equilibrium ("projection"). Thus the delay "becomes the cradle of 'conscious experience'. . . . The motivating force which propels an idea to consciousness is thus conceived of . . . as its drive-cathexis. . . . Ideas may be considered safety valves of drive-tension even if they dispose only of relatively small amounts of cathexes" (p. 691). The drive-cathected memory trace becomes a *drive-representation*, an idea, in the form of consciousness that is characteristic of the primary process. This characteristic is that memories organized around a drive are not clearly differentiated. "The memory of any segment or aspect of experience accrued in the periods of delay, and around the gratification, may emerge as an indicator of mounting drive-tension" (p. 693). In the primary process, therefore, "the cathectic energy in a drive-organization of memory can freely move and center on any representation" (p. 693). The interchangeability of representations of a drive stems here again from "the unrestrained tendency towards discharge by the shortest path" (p. 694) (wish fulfillment). This "mobility of cathexis," free to engage many representations of a drive, manifests itself in the familiar phenomena of displacement, condensation, substitution, *pars pro toto*, and omnipotence of thought. All these thought characteristics of the primary process are then explicitly linked to the

characterizations which Werner and Piaget have used to describe the earlier developmental stages, e.g., diffuse and syncretic thought, undifferentiated structuring of objects and ideas, egocentrism and autism.

2) Secondary process. A twofold differentiation occurs. Drives differentiate into partial drives, and correspondingly into more articulate and discrete representations (ideas). This increasing specificity of drive objects goes hand in hand with a growing differentiation in perception and memory. The memory trace of the drive object now also includes the more clearly experienced *pathways* encountered in finding the need-satisfying object in reality. Therefore the drive representations become less interchangeable, i.e., the cathexes can no longer be displaced to any other drive representation. They depend more closely on *"reality connections."* The cathexes thus limited can no longer press towards direct discharge via hallucination (*"bound cathexes"*). In turn, "memories" replace hallucinatory images. The former short cut of least resistance to wish fulfillment via direct discharge is replaced by a *detour* over object connections in reality (reality testing). These latter attain a psychological representation in the form of *functional* concepts; of action-things; of means-end relations; and of space, time, and causality in the form of naive realism. Ordered, goal-directed thinking now overlays the syncretic forms of thinking. In the further course of development, the functional concepts are also eventually superseded by abstract conceptual ideas and systems of relations. A controlling system, the ego, emerges, superimposed upon the pleasure principle of wish fulfillment which is not, however, replaced. The pleasure principle remains effective, only permitting the detours of reality testing in terms of goal-directed controlled thinking (p. 697).

At those points in this development where conflicts between drive demands and reality, especially delay, become painful, primal repression takes place. It is conceived as the establishment of "countercathexes," similar to the basic means through which the organism originally copes with external stimuli, namely, stimulus-barrier and withdrawal. Energy charges are pitted against the mounting tension of the drive-cathexis, its affect-charge and representation, and energies are withdrawn from the drive-cathexis. In primal repression the "drive and its ideational representation are denied access to motility and consciousness." The energy system of all countercathexes is, of course, conceptualized as censorship and part of the ego structure. "One of the consequences of this structure-development is that not only the repressed drives but all [other] drives are controlled by it. Drive-cathexis alone therefore no longer suffices to raise ideas to consciousness. The nature of consciousness is thus changed" (pp. 695, 700). Drive-cathected ideas which are not repressed, i.e., not countercathected, now have only preconscious quality and need an additional energy to become conscious: attention- or hypercathexis. Thus consciousness becomes a particular form of hypercathecting the preconscious drive representations which have primary process characteristics and can now become organized in terms of the secondary process thought regulations and reality testing.

Up to this point Rapaport seems to have followed Freud closely except for his assimilation of some more recent developmental concepts. His extension of Freudian theory lies, however, in his attempt to integrate the drive-basis of the theory with the autonomous functions of the ego. The higher forms of consciousness, e.g., reflective awareness, forward-looking anticipations and intentions, neutral task-oriented thought processes, and willful creation of new thought patterns and abstractions are all given a relative functional autonomy. Thus he draws on Lewin's experiments on quasi-needs to show how acts of intending may endow objects with valences (cathexes) and create needs in the absence of presently compelling drive-needs. "It is thus that we can think even when no immediate drive-needs motivate us" (p. 704). The integration he attempts cannot here be followed in detail. But one may summarize Rapaport's conception of ego-autonomy as follows: "The ego arises not only from the drive but also from pre-existing apparatuses," the biological givens in the sense of Hartmann. It has energies at its disposal besides the drive energies which it controls. It constitutes a structure of new regulations. These comprise secondary process thought, bound-cathexes, reality testing and "ego-interests." The latter, though arising from drives and *defenses* as derivatives, are nevertheless autonomous and express themselves in quasi-needs, attitudes, wishes, strivings, and values. On the other hand, the ego-structure is vulnerable to the changing cathectic dynamics as in pathology. "It must not be forgotten that, underneath this conceptual-organization there still operates the drive-organization of memories, implying the more archaic forms of concept-formations, such as those we encounter in schizophrenias and in the symbols and mechanisms of dreams" (p. 709).

What are the implications of this model for normal and abnormal thought organization? In essence it amounts to the following thesis: (*a*) There exist different levels and types of cathectic energy distributions; they find expression in drive and thought formation. (*b*) To every level and

type of thought organization, there corresponds a special ego-organization and form or state of consciousness (hypercathexis). As a positive contribution, this may cast a new light on the different stages of mental development in the child where the qualitative differences may be better understood as changing organizations of thought, ego, and consciousness. It also throws new light on schizophrenic thinking in relation to the schizophrenic's form of consciousness and ego-structure. Moreover, in this context, different kinds of consciousness together with their *correlated thought organizations* are postulated. One kind of dimension ranges from the waking stage to daydreaming to dreaming. Within the waking stage, furthermore, different degrees of *awareness* are distinguishable. Maximum awareness is reflective, and is found in the voluntary concentration of intentional thinking of the *normal*. Reduced, reflective awareness is found in the *brain-injured* together with impaired volitional intention in thinking; in the *neurotic*, the reflective awareness is limited (lack of insight); in the *psychotic*, it is either absent or pathologically changed—e.g., to delusions and hallucinations. In each of these states there is a specific organization of thought processes which in turn arises from the interaction between ego-structure, drive-cathexes, and hypercathexes, or, in the language of Rapaport, from the "dynamics of drive-cathexes and their relation to those of cathexes at the disposal of the ego" (p. 730). Rapaport's analysis of attention and concentration has still subtler ramifications than can be here elaborated.

What does this model contribute to our understanding of creative thought and problem solving? How, for example, would Wertheimer's or Duncker's experimental observations be meaningfully supplemented by this view? We are told that in creative activity the ego suspends momentarily its censor function, the "countercathected energy distributions" become momentarily weakened, as also occurs in wit and humor. This leads to the "inventive phase" of creative thinking which "abides by the rule of primary process"—the vague hunch and intuition, for example. Here repressed motives play a specific role. This is in keeping with Kris's thesis that during the creative process, the ego voluntarily regresses into autistic fantasy. Only then does it enter upon the subsequent elaborative phase in which it works out the primary thought pattern in terms of higher-level conceptualization and reality testing. It is noteworthy how creativity is here ultimately linked to the return of the repressed which becomes transformed from the autistic into the realistic-objective. Without making really clear how the content of this repressed material leads to the content of creation, Rapaport does try to connect these contents with the "schematic thought patterns" which stem from earlier drive anticipations (p. 721). One finds it especially hard to conceive how these principles work in productive reasoning as, for example, in the case of young Gauss, or Galileo.

There seems little doubt that from within the psychoanalytic position most of Rapaport's systematizing of Freudian thought on *Thought* will be considered adequate and productive. The reaction of the reviewer who looks "from the outside in" is first the regret that Rapaport's bold theory-construction covers more subject matter than its condensed presentation can account for. It stimulates the reader with a desire for more elaboration of those propositions which Rapaport but briefly states. These include, for example, the origin of drive derivatives, attitudes, and ego-interests. Next, among these and other questions, the following may be focal: is tension discharge, which is the basis of this model, a sufficient description of motivation? Even granted that infant behavior corresponds to this concept, are we justified in making such a process in the *immature* organism the prototype of human adult motivation or deriving it from such an origin? If object-gratification in later stages of maturation and experience show other psychological aspects of goal-achievement, are we still justified in relating these to the inferred tension release on the infant model? This last point is not an original criticism since it has been argued by K. Goldstein and by G. W. Allport [2] as well. In spite of the newer emphasis on "autonomous ego-functions" there remains in psychoanalysis the persistent lack of recognition that already the early stages of child development contain gratifications other than tension discharge. In the interaction between perception, motility, cognition and the object-characteristics as dealt with by the child, the fact that there is more involved than tension release has been shown by the work of C. H. Buehler, W. Stern, Gesell, and others.

It is true that the concept of "drive-reduction" is also central in other contemporary systems such as Hull's and Lewin's. Here analytic ego-psychology could have contributed more psychological sophistication, since motivation is its domain. Rapaport might have followed the lead of Schilder's later work, which advocated that drives, desires, and motives have aims that go beyond the attainment of satisfaction in the sense of the mere restoration of a state of rest; and that one of the basic needs is the drive to go from completeness to incompleteness and then to renewed

[2] Scientific models and human morals. *Psychol., Rev.*, 1947, **54**, 182–262.

completeness. Rapaport might have also drawn upon Piaget's paper in the volume where the physical interaction between organism and environment results in a *change* in the organism, a process Piaget terms "assimilation." This would mean that at least we have to add to the "tension discharge" the modifications and cognitive enrichments that occur in the person during and following the object gratification. One may argue that such changes are provided for in Rapaport's model by such concepts as introjection, identification, and neutralized ego-motives etc. Yet these particular reactions are viewed not so much as the result of object gratification as they are reactions to and defenses against frustration (p. 724 ff.).

This leads to the question of how the relation between the pleasure principle and the reality principle, ego and id, is actually dealt with. It is here where a certain ambivalence in Rapaport's position suggests itself. Though he fully subscribes to and borrows from Hartmann's ego-psychology, he appears in other respects still more bound to orthodox Freudian reasoning than this would lead one to expect. First, he seems impelled to trace the emergence of thought back to *a sequence of experiences centered in the delay of gratification*. The formula: from hallucinatory image to idea to thinking in relations, concrete or abstract, seems to leave more to experience as a cause than to maturational factors. Although the latter are called upon to explain many ego-functions, the development of the control system of the ego poses certain difficulties. We read "The analogy of a river, which where it is slowed down builds up sand bars to slow it further, may help us to visualize what the evidence seems to suggest: the countercathexes seem to be derived from the drive which they repress" (p. 695). A similar problem arises when, on the one hand, Rapaport conceptualizes attitudes, strivings, and values of the ego as ultimately arising from drives as their *derivatives* and, at the same time, attests to their relative autonomy. For example, "the defense of reaction-formation against aggression may manifest itself in the form of 'altruistic' motivation" (p. 701). The ambivalence mentioned above which this reviewer believes he detects may be best illustrated by the following passages, the first from Rapaport, the second from Hartmann, who appears to be more outspoken in regard to the tenets of ego-psychology.

Rapaport: "Our understanding of the process by which the reality-principle arose from the pleasure-principle, the conceptual-organization of memories from their drive-organization, the secondary process from the primary, points the way towards the solution. . . . the original drive-determination of ideation became overlaid by thought-formations which were no longer regulated solely by the wish to reach the gratification-experience by the shortest route, if necessary by hallucination; it was regulated also by the striving to reach it in reality, if necessary by detours. As a result, thinking and action still serve drive-gratification, in final analysis, but at the same time have an independence, autonomy. . . . But however powerful knowledge may grow, and autonomy of thought be extended, the fact remains that one of the ancestors of human knowledge *is* the idea of the need-satisfying object. Knowledge therefore is and will be used directly for the purpose of discharge of drive-cathexes, and not only for experimental action to find the most adaptive path towards gratification" (pp. 722, 723).

Hartmann: "Strictly speaking there is no ego before the differentiation of ego and id, but in the same sense the id does not yet exist either. Both are products of a differentiation process. Those apparatuses which, after this differentiation, are univocally in the service of the ego we will consider inborn ego apparatuses" (p. 396). "Learning to think and learning in general are independent biological functions which exist beside, and in part independently of, drives or defenses" (p. 370). "Besides, the newborn infant is not wholly a creature of drives; he has inborn apparatuses (perceptual and protective mechanisms) which function purposively. . . . A state of adaptation exists before processes of adaptation begin to function" (p. 384). "We know that the reality-principle is, in a sense, a pursuit of the pleasure principle by other means. But this step—the ability to renounce immediate pleasure-gain in order to secure a greater one in the future—cannot be derived from the pleasure-principle alone; not even memories of pain experiences suffice to explain it. We are, however, familiar with the function of anticipating the future and orienting our actions according to it (and with the function of relating means and ends correctly to each other). It is a function of the ego and certainly an adaptation-process of the highest significance. We are entitled to assume that ego-development enters this process [of accepting the reality principle] as an independent variable, though naturally the ego-function in question here may secondarily yield pleasure" (p. 382).

The contrast between these two statements, in terms of the printed word at least, suggests that the first author is still influenced by the "method from below"—namely, the tendency to derive higher levels of organization (thought) from lower and earlier ones (drive delay); whereas the second author accepts the emerging autonomy of ego-functions as generally independent of these "lower" levels. Such differences in approach may lead to important differences in the understanding of certain pathological conditions. This can be illustrated by the four cases of amnesia with organic brain damage, studied by Buerger-Prinz and Kaila. Based on a number of ingeniously-designed test situations, their findings were: inability to do several things simultaneously; alteration of Gestalt function manifest in dedifferentiation, fragmentation and vagueness of intentional goals; disturbance of synthesis of thinking and imagining; premature satiation of the need to find meaning; dependence on and living in the present situation; inability to grasp relationships which lie outside of the momentary situation; inability to relate a present content to total life experience;

inability to assume an attitude by oneself—the attitude has to be induced by outside influences; perseveration and rigidity; lack of self-reflective awareness and self-correction. "As he (the patient) improves, the impersonal qualities of his experience abate" (p. 679).

Buerger-Prinz and Kaila claim that the basis of this disorder *"seems to be the defect in the personal sphere, the passivity and paralysis of the vital layer"* (p. 684). Rapaport comments at this point, that "impulses, drives, and affects are inert. . . . A new impulse sweeps away all previous thought, apparently because there it not energy available for both" (p. 680). The above findings, however, have striking similarities to those which Goldstein and his co-workers discovered, their conclusion being that the basic disturbance is an impairment of *abstract attitude*. On the other hand, in Buerger-Prinz and Kaila's carefully written paper, there is practically no evidence reported for a reduction of drive-energy which would cause this disturbance ("from below"). One wonders whether another more obvious interpretation could be equally valid, namely, that the impairment of the higher cognitive functions is the cause of the symptoms of "the passivity and paralysis of the vital layer"; and that the improvement of those cognitive functions replaces the "impersonal quality of the experience" by "the reappearance of the personal experience structure." With this example, it is *not* suggested that the interpretations of other mental pathology such as psychoses and neuroses should be based on cognitive dysfunctions without consideration of motivational dynamics. However, following the position of analytical ego-psychology, one may pose the following question. It is alleged that in mental disease there is a change of the thought organization, along with changes in the state of consciousness and ego-organization. Is it not also conceivable, perhaps, that the motivational state, too, becomes *reorganized in a new way*? And, in this way, the phenomenon would not be a revival of the archaic drives of the past, disrupting a "new order," but rather a reorganization on all levels, including the motivational structure. It might seem that ego-psychology is almost at the threshold of such a "Copernican revolution," as it were, or is this thought just a "Freudian wish"?

MARTIN SCHEERER

University of Kansas.

AN INTRODUCTION TO PSYCHOLOGY. By Gardner Murphy with the assistance of Herbert Spohn. New York: Harper, 1951. Pp. xx+583. $4.25.

This, the third introductory textbook Gardner Murphy has written, shows a master's touch. The style is direct, clear, and fluent, the illustrations (six in color) are many and vivid, the coverage is wide and the over-all organization clean and transparent. The writing is adorned by references to *belles lettres* (Baudelaire, Gorky, Housman, Shelley, Whitman, Wolfe). The personalistic orientation expresses itself in the many case studies cited and in an emphasis upon individual differences. The index is full, and a useful glossary is provided.

Yet there is a certain lack of discipline which mars the book. Within chapters, the organization is amorphous, indeed sometimes quite confused. The material on imaging (pp. 273 ff.) and on statistics, always a problem to be sure in the introductory course (pp. 374 ff.), seemed to the reviewer especially chaotic. There are *ex cathedra* pronouncements (creation requires a mood—pp. 316–317; prodigies need stimulation—pp. 323–324; intelligence is the ability to learn—p. 333), there is ax-grinding (Burks' data are construed to show that a "considerable" effect [6–9 points] on "intelligence quotient" can be wrought by a definitely favorable environment—p. 350), there are uncritically kind words for ESP, the "subliminal uprush" of Frederic Myers, and Zimmerman's abortive experiments on raising the I.Q. through diet. There are carelessly phrased statements (the US is the stimulus which arouses a response prior to all learning—p. 213; we forget nothing during sleep—p. 264; Burks' method is too complex to be described—p. 350). "Set" is worked to death: it helps explain drive (p. 68), perceiving (p. 146), esthetic preferences (p. 168), learning (p. 219), recall (pp. 258, 262), imaging (p. 273), concept-formation (p. 292), problem-solving (p. 294), creation (p. 312), and intelligence (p. 336). There are non sequiturs: learning is a tension-reduction process, for as learning proceeds there is a gradual reduction of errors (p. 208); sensitiveness is essential to creativeness, for children differ in their sensitivity to color (p. 308); everybody has something of the genius in him, for even geniuses (e.g. Housman) have to work hard (p. 318).

The "suggested readings" at the end of each chapter are inappropriate for the freshmen and sophomores to whom the book is presumably addressed (e.g., Boring's *Sensation and Perception*, Morgan & Stellar's *Physiological Psychology*, Murchison's *Handbook of General Experimental Psychology*, Thurstone's *Factorial Studies of Intelligence*).

These defects, though serious, are not overwhelming. The book remains one with which its intended readers will enjoy working and from which on the whole they can be expected to learn many important and useful principles and facts of psychology.

HARRY RUJA

San Diego State College

INTRODUCTION TO SOCIAL PSYCHOLOGY. By E. Llewellyn Queener. New York: William Sloane Associates, 1951. Pp. 493+viii. $4.50.

The challenge of developing a broad approach to social behavior has been taken up again. Social psychology, ". . . the study of the individual's behavior in response to stimuli from other individuals," is here dealt with through a family of eight basic variables and the four systems of theory considered most adequate for detailing their operation. An interesting book supplied with clear orienting definitions, using ample illustrative content, and written in a simple style is the result.

The idea behind the approach is applied in the following way. Having identified the responses about which he wants better understanding or more effective control, the investigator first studies them in terms of the broadest of the eight variables, the *human-wide* variable. He asks, "if we know only that an individual is a human being, can we predict these particular responses on the basis of that knowledge alone." If so, the introduction of planned changes will result in the controlled occurrence of the behavior under consideration. If, however, the responses in question turn out not to be governed by this dimension, they are then examined from the standpoint of the next narrower one, the *culture-wide* variable. The same procedure is continued through the successively narrower *class, caste, sex, group, crisis,* and the *individual-wide* dimensions until that level of variable (or combination of them) in terms of which these particular responses can be predicted is established. Having achieved this point, the conditions under which the variables will produce the responses (or the conditions under which the responses can be modified) are analyzed to the point where theory no longer provides a basis for understanding and prediction.

The approach is used in a slightly different way in writing the book. Given the *human-wide* variable and the theory of its operation, for example, what kinds of responses can be predicted? Illustrative problems examined in this manner include self-preservation and the so-called maternal, competitive, and heterosexual instincts. Queener concludes that a conception of human nature which provides a fixed, inflexible model for basic social actions is in no way supported by the available evidence. Acceptance of the alternative results in a strong emphasis on learning theory throughout the book.

The treatment of the *culture-wide* variable is wisely limited to a few sample areas. Attention is given to the problem of values and changes in value orientation corresponding with cultural variations in time and place. This chapter, one of the best in the book, is built on a variety of detail ranging from the value orientation of the Roman Empire to that of the West in the various stages of the Christian era. It also includes an interesting comparison of the Soviet and our own value-orientations. The general significance of the *culture-wide* variable is nicely summarized in the statement that, "Although a knowledge of culture gives us the major premises behind an individual's behaviors, this knowledge alone is seldom adequate for predicting specific behaviors. . . . we must know also the minor premises provided by his class, caste, role, situation and personality. . . ."

The *class-wide* variable is introduced through an analysis of class reactions to such values as earning a living, making a home, and training the young. *Caste-wide* behaviors are dealt with through a penetrating use of the literature concerning race differences. The concept of role comes into its own in chapters dealing with the *sex-wide* and the *group-wide* variables. The latter also provides a fairly extensive analysis of leadership behavior.

A somewhat different orientation is involved in the definition and analysis of the *crisis-wide* variable. It includes ". . . any set of conditions presumed by members of a given group as requiring or permitting extraordinary responses." Illustrative material identifies frustration, power, competition, and unstructured situations as ones which provoke *crisis-wide* behavior. Queener's discussion of behavior in an unstructured context is particularly good.

The last dimension introduced to characterize social behavior is the *individual-wide* variable. One chapter emphasizes personality structure and the other one concentrates on social attitudes. The first of these points to the general significance of the *individual-wide* variable for social psychology, rather than providing a conceptual framework which would make the interdependence secure. Neither Freud nor fantasy come into the picture here, and although the idea of *basic personality* is introduced, the treatment of personality organization remains at surface level. For example, while value orientation is one of the major topics discussed in connection with the *culture-wide* and *class-wide* variables, the concept of *superego,* which might well have been introduced to interconnect them with the *individual-wide* variable, is not mentioned in this chapter. Furthermore, the focal concepts of interaction and interpersonal behavior are omitted as are the work of such men as Mead and Sullivan. The chapter on social attitudes moves in the direction of relating cognitive structure and motivational processes to one another. Here again the presentation might have been broadened but the literature is skillfully used to point up the significance of the social determinants of attitude formation and attitude change.

After a discussion of applications in the clinical and industrial fields, two final chapters, one on *theory* and the other on *methodology* are provided. The *theory* chapter, which is located next to the final one in the book, would have made an excellent introductory Chapter II, serving as a general orientation to the uses made of theory throughout the book.

The volume is written for those who want a broad introduction to the psychology of social behavior. Specific behaviors are introduced and occasionally discussed in considerable detail, but the book is properly concerned with showing the operation of this system of analysis, rather than with cataloguing the details of social psychology. The breadth of the foundation on which the approach is based requires the weighing of a very large literature of very uneven quality. The author has done an excellent job of it.

ROBERT R. BLAKE

University of Texas.

PERSONALITY. By David C. McClelland. New York: William Sloan Associates, 1951. Pp. xvi+654. $5.50.

In his preface, McClelland presents his book as "a theoretically oriented text in the psychology of personality"; he adds also that it is "intended as a contribution to the theory of personality." He describes it further, and accurately, as an advanced text presupposing prior work in psychology, and as eclectic, not bound to any one school. Nevertheless, the book's theoretical position is located close to that of the Yale learning theorists.

Taking it first as a text, there is a good deal to recommend the book. It is written in a clear, readable style, very attractively designed and printed, with a minimum of printer's errors; it is up to date, most of the references being to work published since Allport's standard text. The "Notes and Queries" after each chapter are on the whole unusually good; in them the student is again and again encouraged to think for himself, to challenge statements in the text, and to design experimental ways of testing and verifying the ideas that are presented. Finally, the presentation is unified around a single concrete case, Karl, about whom quite a lot of data are presented. In addition, there is an instructor's manual that contains all of the complete protocols on Karl, including some crucial data not found in the book—answers to predictive questions which the student is asked to fill out at various points in the presentation.

Nevertheless, I cannot recommend the book as the basic text for a course in personality. Despite its many merits, it would mislead and confuse students on too many important issues. Most of the reasons for this judgment will be presented below, in a discussion of the theory. In addition, it has the fault of encouraging a too-simple view of man's complexities, an attitude that the apparent complications are due to faulty theorizing rather than the nature of the subject matter itself. The essential superficiality of most of the data about the illustrative case and the omission of crucial aspects of his history contribute further to this impression.

The same unevenness is apparent when we consider the book as a contribution to theory. Stimulating original ideas are found side by side with sloppy, poorly thought through reasoning, and irresponsible speculations blandly presented as fact. Time and again I found myself marking passages to come back to for their balanced, insightful critique (for example, discussions of the nomothetic-idiographic pseudo-issue, the method of matching, the problem of validation in personality research, and factor analysis). But even more often I had to note careless thinking, evasion of important issues (cf. the discussion of similarity, pp. 220–224), inconsistency in the critical standards that are applied to different theories, naive speculations about schizophrenia (cf. pp. 200, 509, 540) and other complex clinical conditions with which the author has apparently had no direct contact, and so on. It is regrettable that McClelland didn't let his book ripen a few more years before bringing it out; he is clearly capable of doing much better.

The first of the book's five major divisions, on methodology, gets off to a good start with a beautiful introductory chapter that sets the stage for a scientific consideration of personality by giving a brief but thought-provoking review of its background in intellectual and cultural history. The rest of the section is a distinct letdown. It asks a good many of the right questions but fails to provide clear and consistent answers to them.

One of the ways in which McClelland achieves the narrow canvas with which he finds himself comfortable in portraying personality is by ruling out large areas on "methodological" grounds. His first criterion for an acceptable concept is that it must "function according to the principles of learning," because "we know a great deal about the way in which a person learns." The meaning of this odd stipulation is that no concepts are allowable to deal with innate or hereditary aspects of personality! Thus the study of personality is arbitrarily restricted to what is learned, and the implicit assumption creeps in that there *is* nothing innate, that all determinants of behavior are

learned. No better way exists to cripple the growth of a science than to restrict its concepts to what is best known at any point in time.

The second main principle of exclusion is to define as outside the realm of science all *evaluative* concepts. This is a tricky point. McClelland is quite right, following Allport, in excluding moralistic evaluations (e.g., good, sinful) from the tools we use to describe or understand personality. Such terms tell us almost nothing about the person and change the character of our approach to him. But McClelland would exclude much more: concepts like *intelligence* or any other ability or aptitude are ruled out, likewise descriptions of the *S*'s relation to a social norm, like *law-abiding*. But objective operations can be easily specified for concepts like these; they need not have any normative tinge. To turn one's eyes away from the problem areas that they encompass for fear of appearing to pass value judgments is to impoverish personology in a quite unnecessary way.

The next three major sections in turn take up *trait*, *schema*, and *motive* as personality variables; these, with the addition of a derivative one, the self-schema, are McClelland's main conceptual tools and define the scope of personology as he sees it. (The theory has no place for psychosomatics, for problems of thought organization or psychopathology generally.) In the final, integrative section, McClelland says (p. 611) "our theoretical position is that traits, schemata and motives enter into the determination of any act." Yet most of the time he talks as if a particular act were the effect of one or the other of these (cf. queries on pp. 331, 619). This inconsistency between general statement of theoretical position and actual practice is unfortunately met with very frequently in the book. The final "synthesis" is weak in principles that pull the parts together, so that the theory as it stands would have difficulty in accounting for the unity and integration of many personalities—as it does for Karl's.

I found his approach to definitions generally a strange one. By defining personality not as the subject matter to be conceptualized by personology but as the conceptualization itself, he neatly avoids the issues that have traditionally been wrestled with in the 50 definitions Allport presents. And he simply fails ever to define the crucial (to an S-R theory like this) concepts *stimulus* and *response*. Then he puts into the definition of the concepts *trait* and *motive* that they are learned, later speaking about the acquired nature of traits as if it had been proved, not just assumed. It is dangerous and unnecessary to introduce an hypothesis growing from a theoretical bias into a definition—especially when a term like trait is used as loosely as McClelland likes to do. For he speaks of "physique traits" such as the somatotype; is the somatotype learned, then? Sheldon's untamed geneticism needs an antidote, but not such an extreme one as this.

The concept *schema* is used to cover all structures and ideational contents: information, percepts, memories, attitudes, etc. Such a concept may be useful in unifying this diverse field; it draws together such different work as Sherif's experiments and theories about social norms, Gestalt psychological contributions on the development of stable memory traces, and the social psychology of attitudes. This section contains a discussion of socialization in which good use is made of cross-cultural materials and learning principles. For it is through the formation of the individual's schemata that McClelland sees the main impact of culture on personality.

The section on motivation contains a fair amount of material, mostly distorted, on psychoanalysis. It is sad to have to report that the treatment of psychoanalysis throughout this book is marred by the same ignorance, condescension, straw-man arguments and naive imputations that have been traditional in academic texts. The author talks freely about what "the psychoanalysts say" as if he had an extensive knowledge of the psychoanalytic literature, but constantly displays his ignorance by statements such as that Freud developed "his theory of the Oedipus complex from his study of Little Hans." The critique of other motivational theories is somewhat better, but it is not until the chapter on experimental studies of motivation that McClelland comes into his own. In it he builds up a very convincing case for what to me is the main theoretical contribution of the book—the hypothesis that affective learning in childhood is so persistent and "hard to unlearn because the learning is so general in the first place, so compounded of different cues, responses, rewards, and punishments, that it will be hard for the person ever to discover that conditions have changed, that some general expectation he has formed is no longer being confirmed" (pp. 442–443). He reports an ingenious experiment with which he has tested and verified this theory.

In large part, the book is organized around the interesting set of researches the author and his collaborators have done on n Achievement. Despite the cleverness, persistence, and interrelatedness of this work, it has a serious weakness: the arbitrariness with which the decision was made in the first place, with no explicit criteria and in the face of internal evidence in his research (not reported in this book) *against* the interpretation, that the

phenomenon measured by his method of scoring the TAT *was* n Achievement. Furthermore, he contends that his is the best available measure of this need, because the influence of stimulating conditions and of schemata are assumed to be at a minimum in imaginative products. This assumption is not only unproved but unlikely; there is plenty of evidence against it.

One final criticism of this often stimulating, often exasperating book. Again and again, the author falls into a trap which he makes for himself out of too great a respect for the external trappings of experimental method, and too little appreciation of clinical observation as a source of data. There are two stages to this error. In the first place, he over-values experiment; if a few observations of a few subjects have been made in the context of an experimental setting, he is far more impressed with them and relies on them far more than the reports of no matter how many observations of behavior in clinical settings. Experimental controls *are* extremely important, and we need much more of them in personology; but they cannot substitute for *relevance*. Then in the second place, McClelland seems to feel so secure when he has a finding from a controlled research design to build on that he often forgets all caution, all the rules of scientific thinking, in generalizing these findings and making all kinds of loose derivations and extensions from them. Thus he falls into the same kind of unscientific thinking of which he correctly accuses Freud. Many examples could be cited; here is just one. Whiting and Child found in a study of child-rearing practices and theories of disease in over 50 cultures that explanations of disease as due to influence of spirits were significantly associated with "either high nurturance or neglect in infancy." McClelland therefore suggests, as a possible explanation of the fact that Karl sometimes hears voices, the fact that his parents rejected him.

In the book's first pages, the author makes the following unusual confession: "I have never been psychoanalyzed, never been to a Rorschach Institute, cannot properly interpret a Thematic Apperception Test, space 'mhms' correctly in a nondirective interview, deduce a theorem from Hull's postulates, or ask a Navaho informant about his sex life. But perhaps all this is what qualifies me to write this book" (without special pleading for any school). Perhaps also, the reader may add, it would have been a better book if the author had been less uncontaminated by these kinds of direct experience with his subject matter.

ROBERT R. HOLT

The Menninger Foundation,
Topeka, Kansas.

THE ADOLESCENT PERIOD: A GRAPHIC ATLAS. "THE ADOLESCENT PERIOD: A PICTORIAL ATLAS. MONOGR. SOC. RES. CHILD. DEVELOPM., 1949, 14 (1, 2). Edited by Frank K. Shuttleworth. Evanston, Ill.: Child Development Publications, 1951. Pp. 273, v. 69.

These two companion monographs have been organized in order to present the data now available on the topics of physical growth, sexual maturity, and body builds. They have been designed to be "useful as supplementary material for courses in psychology, child development, education, physical education, sociology, public health, nursing education, and pediatrics." The graphic data and the photographs "have been assembled in separate monographs partly as a matter of convenient publication and partly to permit the more flexible use of the materials."

The "graphic atlas," which is a revision of an earlier monograph published in 1938, contains 453 figures. These figures are organized in 13 topical chapters which tell a large part of the story about (1) adolescence and the adolescent population; (2) physical (bodily) growth; (3) physical (bodily and to some extent behavioral) growth and differentiation; (4) physiological functions; (5) sexual maturation (the largest chapter with 76 figures); (6) health; (7) intelligence; (8) education; (9) special abilities; plays and games, interests, attitudes, and problems; (10) behavior maladjustments; (11) occupational adjustments; (12) sex adjustments (containing nine figures from the Kinsey report), and (13) education, occupation, and marriage. Although verbal text is limited to brief explanatory legends for these figures, the fact that each legend contains a reference to the original publication from which it was drawn makes this a very useful anthology for graduate students and mature scholars from other fields. Undergraduate lecturers should find it a useful supplementary assignment and a good source of slides with which to illustrate lectures.

The "pictorial atlas" contains 400 original nude photographs arranged in 52 figures to show both developmental changes within the same individuals and the differences between individuals. The individual differences depicted concern both the longitudinal contrasts between early and late developers and cross-sectional differences between individuals in body build and sexual characteristics. The longitudinal comparisons and contrasts are especially well shown by pictures of given individuals taken annually over spans of age as long as from three to 18 years. Cross-sectional comparisons are illuminated by presenting a variety of individuals at a given age. For example, Chapter 7 is devoted to "variations in body builds

at age fifteen." Sheldon's somatotypes receive two chapters, one devoted to variety and to change with age, and the other to the effects of partial starvation. The tenth and final chapter gives pictorial illustrations of a series of androgeny ratings. The text in this pictorial atlas is limited to brief comments which indicate whether the depicted relationships are typical, universal, or relatively unusual. In this reviewer's opinion, this pictorial atlas should be highly useful to teachers of the courses for which Dr. Shuttleworth prepared it.

These two monographs constitute a distinct contribution to our battery of teaching materials on the topics of physical growth and development.

J. McV. HUNT

University of Illinois.

STUDIES IN LEADERSHIP: LEADERSHIP AND DEMOCRATIC ACTION. Edited by Alvin W. Gouldner, New York: Harper and Brothers, 1950. Pp. xvi+736. $5.00.

In compiling the 31 essays (20 of them previously unpublished) composing this volume, Gouldner has drawn upon the resources of sociology, political science, social philosophy, and social psychology. Five of the 34 authors respresented are social psychologists, strictly defined, and surely the majority of the points of view in the book are sociological and historical.

Gouldner's main aim in preparing this book seems to have been to bring social science scholarship into closer rapport with social action, particularly in democratic action groups. The editor does not attempt to hide his advocacy of democratic leadership and group organization, nor is this alliance with democracy merely an ethical, apolitical one. He states his point of view clearly when he says (p. 14) ". . . an effort has been made to focus on those problems of leadership which seem specifically helpful as guides to democratic action or which, more generally, round out the perspective of an individual and promise aid in making him a more informed *participant* in a democracy." Most of the essays included in the volume sustain this general view.

In addition to an Introduction, the editor has written brief commentaries for each section that serve to set the stage for the individual papers. These "Contexts," as Gouldner calls them, place the ideas of each contributor in a line of historical development, show his relation to a particular systematic position, or interconnect the several essays in one section. The range of material covered is indicated by the titles of the five major parts into which the papers are grouped: Types of Leaders; Leadership and Its Group Settings; Authoritarian and Democratic Leaders; The Ethics and Techniques of Leadership; and, finally, Affirmations and Resolutions—four "summary" papers that point up the diversity rather than the homogeneity of social scientists' thinking about leadership.

Such a summary well reflects the nature of the whole volume. Gouldner's most valuable contribution perhaps is to shatter any remaining illusions a reader may have that "leadership" is a unitary phenomenon or that any one social science discipline has prior rights or superior analytical schemata for handling leadership problems. Leadership is like the weather—everybody talks about it.

Gouldner has not tried to minimize the disparity that exists in leadership research. His book is not a comprehensive text nor does he propose a systematic theory of leadership. Rather, in his own final view, "we may find unity in diversity," of points of view. This conclusion will bring cold comfort to one who is looking for a compact, uncomplicated fomulation of leadership problems; for, from such a point of view, Gouldner's book leaves the job half done. The task of presenting different points of view is accomplished although social psychologists are likely to feel that quantitative and experimental research has been neglected in favor of observational and philosophical-theoretical work.

But the more difficult task of ordering and systematizing points of view is just begun in the editor's Introduction. In these 46 pages, Gouldner considers briefly the factors that have aroused interest in leadership as a social and a scientific problem and then goes on to an apt examination of the "trait" and "situationist" approaches to its study. This discussion draws its terms and directions from two main schools of social science thought: the sociological structural-functional, represented by Parsons and Merton; and the social psychoanalytic—Fromm and Redl. While the argument is conducted in abstract terms and little reference is made to concrete data, social psychologists will find many critical insights that can redirect their formulation of empirical problems, particularly in Gouldner's examination of the universality of both character traits and situational features in leadership.

At the same time, psychologists may enjoy having their horizon of research broadened by considering Gouldner's suggestive argument that the very nature of leadership research and the terms in which problems are stated is influenced by the kind of society in which the researcher lives and by the sources of support for his investigation. Bureaucratic organizations for example, have supported much leadership research; and, because of the nature of a leader's activity in a bureaucracy, such research has fostered the "job-analysis" view of leadership. Gouldner goes on to point out a

secondary consequence of such formulations—that social psychologists have tended to separate leadership from authority (deriving from office) and tended to neglect the latter order of problems.

But while he may catch glimpses of new terrain, the social psychologist will not find his own backyard well surveyed, since only four papers in the collection belong there. Most of Lewin's (1939) paper reprinted as "Experiments in Social Space" in *Resolving Social Conflicts* appears in Gouldner as "The Consequences of an Authoritarian and Democratic Leadership." Doob summarizes some old and some recent psychological research in a hitherto unpublished paper entitled "Propagandists vs. Propagandees." Kutner's previously unpublished essay "Elements and Problems of Democratic Leadership" is less a research report than a summary of a democratic point of view on leadership activities. A ten page excerpt from *The People's Choice* includes both data and inferences drawn by Lazarsfeld, Berelson, and Gaudet on "Informal Opinion Leaders and a National Election."

Individual critical treatment of the remaining essays is clearly impossible within the scope of this review, but it seems appropriate to comment on some of them, especially those that both are of more than incidental interest to social psychologists and have not been published previously. T. W. Adorno's "Democratic Leadership and Mass Manipulation" analyzes the relationship between character structure and techniques of public persuasion. Bernard Barber's treatment of "Participation and Mass Apathy in Associations" provides both empirical data and insights into an aspect of democratic group activity often ignored by psychologists. Barber's demographic and social-structural emphasis is supplemented by Riesman and Glazer's historical and psychological orientation in their discussion of "Criteria for Political Apathy," based on intensive personal interviews.

That this review fails to comment specifically on many of the remaining papers does not mean that they are without value or significance to social psychology. Most of them treat leadership problems occurring in larger and more complex social organizations than the experimental social psychologist is accustomed to study. Gross, macroscopic factors (from the laboratory point of view) in leader behavior take the center of the stage and bring with them a sociological-historical rather than a questionnaire-interaction scoring technique of data gathering and presentation. A rigorous, quantitatively-minded social psychologist may be made uneasy by such essays as Lowenthal and Guterman's quotation-packed "Self-Portrait of the Fascist Agitator," Norman Miller's lively "The Jewish Leadership of Lakeport," or Whyte's shrewd and sensitive analyses of informal leaders' behavior in Cornerville, but he cannot help being enlightened and made more aware of the complexity of leadership in concrete situations.

By way of a summary, this reviewer sees Gouldner's book as having several possible uses. First, its Introduction should stimulate the reformulation of some leadership problems in social psychology and make psychological researchers more sensitive to social-structural factors. Secondly, many of the essays will be useful reading assignments in courses, especially those given in a "human relations" or "social relations" context. Last, the book contains much concrete descriptive material about unions, communities, business organizations, and other groups which will be of considerable assistance in orienting social psychologists to the problems they must face if they attempt leadership research in non-laboratory situations.

HENRY W. RIECKEN

Harvard University.

NEW BOOKS RECEIVED

ANDERS (STERN), GUNTHER. *Kafka, Pro and Contra.* Munchen, Germany: C. H. Beck'sche Verlagsbuchhandlung, 1951. Pp. 109.

ANDERSON, HAROLD H., AND ANDERSON, GLADYS L. *An introduction to projective techniques.* New York: Prentice-Hall, 1951. Pp. xxiv+720. Price $6.75.

BASH, K. W. *Consciousness and the unconscious in depth and Gestalt psychology.* Amsterdam, Holland: North-Holland Publishing Co., 1949. Pp. 83.

BUELL, BRADLEY, AND ASSOCIATES. *Community planning for human services.* New York: Columbia Univ. Press, 1952. Pp. xiv+464. Price $5.50.

BUNDESEN, HERMAN, N., M.D. *Toward manhood.* New York: Lippincott, 1952. Pp. 175. Price $2.95.

CARRINGTON, HEREWARD, AND FODOR, NANDOR. *Haunted people, Story of the Poltergeist.* New York: E. P. Dutton, 1951. Pp. 222. Price $3.50.

CLEUGH, M. F. *Psychology in the service of the school.* New York: Philosophical Library, 1951. Pp. vii+183. Price $3.75.

CORY, DONALD W. *The homosexual in America.* New York: Greenberg, 1951. Pp. xvii+326. Price $4.00.

FIRTH, RAYMOND, F. B. A. *Elements of social organization.* New York: Philosophical Library, 1951. Pp. vii+257. Price $5.75.

FLESCHER, JOACHIM, M.D. *Mental health and the prevention of neurosis.* New York: Liveright, 1951. Pp. xi+603. Price $5.95.

FLÜGEL, J. C. *A hundred years of psychology.* (2nd edition). London: Gerald Duckworth and Co., 1951. Pp. 424. Price $3.25.

FODOR, NANDOR. *New approaches to dream interpretation.* New York: Citadel Press, 1951. Pp. xvi+368. Price $5.00.

FREUD, SIGMUND. *Psychopathology of everyday life.* New York: Mentor Books, 1951. Pp. 168. Price 35 cents.

FREUD, SIGMUND (Trans. Strachey, James). *Totem and taboo.* New York: W. W. Norton, 1950. Pp. ix+172. Price $3.00.

GINDES, BERNARD C. *New concepts of hypnosis.* New York: The Julian Press, 1951. Pp. xv +262. Price $4.00.

GOVERNOR'S STUDY COMMISSION. *Report on the deviated criminal sex offender.* State of Michigan, 1951. Pp. xi+245.

HARTWELL, SAMUEL W., M.D. *Sexual abnormalities.* Wash., D. C.: Public Affairs Press, 1951. Pp. viii+71. Price $1.00.

HOLLIS, ERNEST V., AND TAYLOR, ALICE L. *Social work education in the United States.* New York: Columbia Univ. Press, 1951. Pp. xviii+ 422. Price $5.50.

JOSEPH, ALICE, M.D., AND MURRAY, VERONICA F., M.D. *Chamorros and Carolinians of Saipan.* Cambridge: Harvard Univ. Press, 1951. Pp. xi +381. Price $5.00.

KELLY, E. LOWELL, AND FISKE, DONALD W. *The prediction of performance in clinical psychology.* Ann Arbor: Univ. of Michigan Press, 1951. Pp. xv+311. Price $5.00.

KUHLEN, RAYMOND G. *The psychology of adolescent development.* New York: Harpers, 1951. Pp. xvii+675. Price $5.00.

MEAD, MARGARET, AND MACGREGOR, FRANCES COOKE. *Growth and culture, A photographic study of Balinese childhood.* New York: G. P. Putnam's Sons. 1951. Pp. xvi+223. Price $7.50.

MILLER, GEORGE A. *Language and communication.* New York: McGraw-Hill, 1951. Pp. xiii+298. Price $3.75.

MOLONEY, JAMES CLARK, M.D. *The battle for mental health.* New York: Philosophical Library, 1952. Pp. x+105. Price $3.50.

NATIONAL CONFERENCE OF SOCIAL WORK. *The social welfare forum, 1951.* New York: Columbia Univ. Press, 1951. Pp. xvi+380. Price $5.00.

NATIONAL OPINION RESEARCH CENTER. *The role of unemployment compensation in maintaining family income and expenditure in an area of critical unemployment.* Chicago: National Opinion Research Center, 1951. Pp. iv+37.

PARKER, WILLIAM R. *Pathology of speech.* New York: Prentice-Hall, 1951. Pp. xiv+321. Price $3.25.

PARSONS, TALCOTT. *The social system.* Glencoe: Free Press, 1951. Pp. xviii+575. Price $6.00.

PARSONS, TALCOTT, AND SHILS, EDWARD A. (Eds.) *Toward a general theory of action.* Cambridge: Harvard Univ. Press, 1951. Pp. xi+ 506. Price $7.50.

PLOSCOWE, MORRIS. *Sex and the law.* New York: Prentice-Hall, 1951. Pp. ix+310. Price $3.95.

RALLI, ELAINE P. (Ed.) *Adrenal cortex* (Transactions of 2nd Conference, 1950, sponsored by Josiah Macy, Jr. Foundation). New Jersey: Progress Associates, Inc., 1951. Pp. 209. Price $3.00.

REDL, FRITZ, AND WINEMAN, DAVID. *Children who hate.* Glencoe: Free Press, 1951. Pp. 253. Price $3.50.

SEATON, GEORGE JOHN. *Isle of the damned.* New York: Farrar, Straus and Young, 1951. Pp. xvi+302. Price $3.50.

SENN, MILTON J. E., M.D. (Ed.) *Problems of infancy and childhood* and *Supplement I—Family centered maternity and infant care.* (Transactions of the 4th Conference, 1950, sponsored by the Josiah Macy, Jr. Foundation.) New Jersey: Progress Associates, Inc., 1951. Pp. 181. Price $2.50.

SENN, MILTON J. E., M.D. (Ed.) *Symposium on the healthy personality.* (Transactions of special meetings of the Conference on Infancy and Childhood, 1950, sponsored by the Josiah Macy, Jr. Foundation.) New Jersey: Progress Associates, 1950. Pp. 298. Price $2.50.

SINGER, ERIC. *The graphologist's alphabet.* New York: Philosophical Library, 1951. Pp. 117. Price $3.75.

STERN (ANDERS), GUNTHER. *Kafka, Pro und Contra.* Munchen, Germany: C. H. Beck'sche Verlagsbuchhandlung, 1951. Pp. 109.

STOOPS, JOHN DASHIELL. *The integrated life.* New York: Richard R. Smith, 1951. Pp. 180. Price $3.00.

TALBERT, ERNEST L. *In the spirit of Jefferson,* Essays and Reviews. New York: Exposition Press, 1951. Pp. 65. Price $2.00.

TINBERGEN, N. *The study of instinct.* New York: Oxford Univ. Press, 1951. Pp. xii+228.

WATSON, ROBERT I. *The clinical method in psychology.* New York: Harpers, 1951. Pp. xii+779. Price $5.00.

WHITE, RALPH K. *Value-analysis, The nature and use of the method.* Society for the Psychological Study of Social Issues, 1951. Pp. 87. Price $1.00.

WILLIAMS, GARDNER. *Humanistic ethics.* New York: Philosophical Library. 1951. Pp. xii+223. Price $3.75.

WORMSER, MARGOT H., AND SELLTIZ, CLAIRE. *How to conduct a community self-survey of civil rights.* New York: Association Press, 1951. Pp. xiv+271. Price $3.75.

ZALEZNIK, A. *Foreman training in a growing enterprise.* Boston: Harvard Graduate School of Business Administration, 1951. Pp. xv+232. Price $3.50.

HISTORY OF AMERICAN PSYCHOLOGY

by A. A. Roback

Here is the first history of American Psychology ever to appear, showing through development stages how this vastly significant aspect of human study reached its present importance. The volume presents an over-all picture covering three centuries, including the numerous divisions and activities of the powerful American Psychological Association.

Author of more than twenty books on human behavior (many translated into foreign languages), and as one who stood close to the chief architects of the science, Dr. Roback naturally possesses much first-hand information. The ever-growing importance of the subject to students, researchers, psychologists, and intelligent laymen renders this an invaluable tool for study, reference, and genuine interest. Copiously illustrated. $6.00

Expedite Shipment by prepayment

LIBRARY PUBLISHERS

8 West 40th St., New York 18, N. Y.

BERTRAND RUSSELL'S DICTIONARY OF MIND MATTER & MORALS

This exhaustive work offers more than 1000 definitions and opinions of the 1950 Nobel Prize winner, arranged as a handy key. Here is Russell's challenging thought on politics, ethics, philosophy of science, epistemology, religion, mathematical philosophy, and on topics crucial to an understanding of international affairs today. Dipped into casually it rewards the browser with stimulating and acute intellectual insights. Read intensively it will be found indispensable to a fuller appreciation of one of the profoundest minds of our age.

$5.00

Expedite Shipment by prepayment

PHILOSOPHICAL LIBRARY

Publishers

15 East 40th Street, Desk 185

New York 16, N. Y.

THE INTERNATIONAL SOCIAL SCIENCE BULLETIN

published quarterly by the United Nations Educational, Scientific and Cultural Organization, 19, Avenue Kléber, Paris, France.

TO acquaint social scientists in various countries with work conducted by national and international organizations, as well as individuals, concerned with an objective approach to the study of human relations.

Specimen copy sent on request

Annual Subscription: $3.50

Single Copies: $1.00

UNESCO

COLUMBIA UNIVERSITY PRESS

2960 Broadway **New York 27**

Pastoral Psychology

A Religiously Oriented Monthly Periodical of the Science of Human Behavior

As Dr. John A. P. Millet, Chief Psychiatrist of the American Rehabilitation Committee, said in a recent letter introducing our journal to the members of the American Psychiatric Association, "Science which is only interested in mechanistic solutions leaves its exponents bewildered and frustrated . . . religion can provide the conviction that the goals of our patients' efforts to get well are worth the struggle . . ."

Dr. William C. Menninger says: "PASTORAL PSYCHOLOGY fills a much felt need. . . . I am confident that it will be of much interest and helpful to psychiatrists."

Here are a few of our important articles:

THE MATURE PERSONALITY	Gordon W. Allport, Ph.D.
RELIGIOUS APPLICATIONS OF PSYCHIATRY	Karl A. Menninger, M.D.
DEALING WITH INTERPERSONAL CONFLICT	Carl R. Rogers, Ph.D.
THE PASTOR AND THE CLINICAL PSYCHOLOGIST	Molly Harrower, Ph.D.
WHAT IS HAPPENING TO THE AMERICAN FAMILY?	Margaret Mead, Ph.D.
RELIGION, PSYCHOTHERAPY, AND THE ACHIEVEMENT OF SELFHOOD	Rollo May, Ph.D.
DISCIPLINE AND MENTAL HEALTH	O. Hobart Mowrer, Ph.D.
THE ART OF DREAM INTERPRETATION	Erich Fromm, Ph.D.
SEX AS AN EXPRESSION OF PERSONALITY AND SOCIAL VALUES	Lawrence K. Frank
THE ACT OF SURRENDER IN THE TREATMENT OF THE ALCOHOLIC	Harry M. Tiebout, M.D.
THE TREATMENT OF THE HOMOSEXUAL	George W. Henry, M.D.
THE GROWTH FACTOR IN CHILD PERSONALITY	Arnold Gesell, M.D.
EMOTIONAL FACTORS OF ACCIDENT PRONENESS	Henry H. Brewster, M.D.
PSYCHOSOMATIC ASPECTS OF LOVE	Gotthard Booth, M.D.
DYNAMICS OF PERSONALITY DEVELOPMENT	Franz Alexander, M.D.
BODY, MIND, AND SPIRIT	John A. P. Millet, M.D.

You, too, will find PASTORAL PSYCHOLOGY a practical, helpful tool in your work.

Dr. Robert H. Felix, Director, National Institute of Mental Health, says: "I have found the articles very pertinent to the work which I am doing and believe the psychiatrist in clinical practice will find them equally significant."

Subscription Form

PASTORAL PSYCHOLOGY
Great Neck, N. Y.

I enclose

☐ $3 for a one-year subscription
☐ $6 for a three-year subscription
(Canadian and foreign subscriptions 50¢ per year additional)

Name.....

Address.....

City..... Zone.... State.....

FUNDAMENTALS OF

Social Psychology

by EUGENE L. HARTLEY *and* RUTH E. HARTLEY

With a sure balance between the psychological and sociological approaches, this text synthesizes the materials of social psychology through the concepts of communication and socialization.

The authors—accomplished social psychologists with broad independent research experience—emphasize the social-developmental rather than the biological-genetic aspects of the subject.

Critical social problems are analyzed by colleagues in several specific areas, and much previously unpublished research is presented. The concepts of functional membership groups, status, role, leadership-followership, and the processes of communication and socialization are fully discussed.

This is an excellent new text for courses in Social Psychology, Social Relations, Communications, and Mass Media.

CONTENTS

1. The Field of Social Psychology
1. Communication—The Basic Social Process
2. The Importance and Nature of Communication
3. Task of the Communicator
4. Task of the Communicant
5. Nature of the Communicant
6. Barriers and Facilitations in Communication
7. Dynamics of Mass Communication (G. D. Wiebe)
11. Socialization—Group and the Individual
8. Socialization is Learning
9. Perceiving
10. Doing and Feeling
11. Further Considerations in Socialization
12. Delinquency (S. L. Sachs)
III. The Individual and the Group
13. Groups and Group Formation
14. Social Norms
15. Individual in the Group
16. Role (I)
17. Role (II)
18. Status
19. Leadership-Followership
20. Attitude and Opinion (with C. W. Hart)
21. Basic Attitudes
22. Modifying Basic Attitudes
Bibliography
Index

6⅝ x 9⅛ inches, cloth, 832 pages, $5.50 net
Just Published

Examination copies on request

ALFRED A. KNOPF, Inc.
501 Madison Avenue, New York 22

BORZOI

New McGRAW-HILL Books

PERSONALITY AND PSYCHOTHERAPY. *An Analysis in Terms of Learning, Thinking and Culture*

By JOHN DOLLARD and NEAL E. MILLER, Yale University. *McGraw-Hill Publications in Psychology.* 479 pages, (*textbook edition*) $6.50.

This distinctive new text deals with abnormal behavior and psychotherapy, since they bring out clearly factors that are crucial in the dynamics of the normal personality. A wide range of material is integrated. Three great traditions are brought together—psychoanalysis, the exact methods of the experimentalists, and modern sociology and anthropology.

PSYCHOLOGY

By ROSS STAGNER, University of Illinois and THEODORE KARWOSKI, Dartmouth College. 582 pages, $5.00.

An outstanding new text for first-year college students, covers the facts of psychology in its important areas (motivation, sensing, perceiving, learning, thinking, intelligence and personality) and presents these facts in such a way that the student can readily grasp both their interrelationships and significance within the total scope of the book itself. In order to effect the exceptional unity of treatment which this book affords, the authors have presented their material in terms of *homeostasis* and the *psychology of perception.*

PERSONALITY MEASUREMENT

By LEONARD W. FERGUSON, Metropolitan Life Insurance Company, New York. *McGraw-Hill Publications in Psychology.* Ready in July.

A college text, providing the first comprehensive account of the major methods by which personality tests have been constructed, this book offers a description and appraisal of the test construction methodologies. The text's concentration on methodology with sufficiently detailed discussions of each method, gives the student a real working knowledge of the basic principles involved.

THE DYNAMICS OF THE COUNSELING PROCESS

By EVERETT L. SHOSTROM, George Pepperdine College, Los Angeles; and LAWRENCE M. BRAMMER, Sacramento State College. Ready in June.

Based on the fundamental concept that the counselor's purpose should not be simply to solve the student's problems, but he should rather aim at facilitating the process of mature self-direction and providing a situation in which the student can help himself. The aim of this new text is to present a unified approach to educational-vocational-personal counseling.

Send for copies on approval

McGRAW-HILL BOOK COMPANY, Inc.

330 West 42nd Street New York 36, N. Y.